

PHYSIQUE SUBATOMIQUE

par

David SÉNÉCHAL

Ph.D., Professeur Titulaire

$$\begin{aligned}\mathcal{L} = & -\frac{1}{4} F_{\mu\nu} F^{\mu\nu} \\ & + i\bar{\Psi} \not{D} \Psi + c.h. \\ & + \bar{\Psi}_i \gamma_{ij} \Psi_j \phi + c.h. \\ & + |D_\mu \phi|^2 - V(\phi)\end{aligned}$$

UNIVERSITÉ DE SHERBROOKE

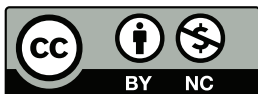
Faculté des sciences

Département de physique

(mars 2020)

Ce manuel électronique fut utilisé dans le cadre du cours PHQ636/PHQ638 (Physique subatomique) à l'Université de Sherbrooke, de 2003 à 2019. Il fait partie d'une collection de manuels électroniques diffusés par des professeurs du département de physique de l'Université de Sherbrooke. Il a été revisité pour une diffusion sous licence libre en collaboration avec la fabriqueREL en mars 2020. Il est diffusé sous licence *Creative Commons* dans sa version BY-NC, sauf indications contraires.

L'auteur, David Sénéchal, est professeur titulaire à l'Université de Sherbrooke. Son domaine de recherche est la modélisation numérique des matériaux quantiques. C'est dans un esprit de partage et de collaboration qu'il a décidé de partager cette ressource éducative libre. La liste de ses publications est disponible sur [Google Scholar](#).



Sauf indications contraires, le contenu de ce manuel électronique est disponible en vertu des termes de la [Licence Creative Commons Attribution - Pas d'utilisation commerciale 4.0 International](#).

Vous êtes encouragé à :

Partager – copier, distribuer et communiquer le matériel par tous moyens et sous tous formats.

Adapter – remixer, transformer et créer à partir du matériel.

Selon les conditions suivantes :

Paternité – Vous devez citer le nom de l'auteur original.

Pas d'utilisation commerciale – Vous n'avez pas le droit d'utiliser le matériel à des fins commerciales.

Figures non conformes à la licence CC-BY-NC :

La figure suivante est reproduite avec permission, mais n'est pas sujette à la licence CC-BY-NC. Il est de la responsabilité du prochain utilisateur d'effectuer les démarches nécessaires auprès du détenteur des droits pour obtenir les permissions :

Fig. 6.3 page 184.

Tous droits réservés pour cette figure.



fabrique REL
RESSOURCES ÉDUCATIVES LIBRES



**UNIVERSITÉ DE
SHERBROOKE**

TABLE DES MATIÈRES

1	Introduction	9
A	Panorama historique	9
B	Le modèle standard : survol	13
1	Classification des particules élémentaires	13
C	Rappels de relativité restreinte	16
1	Transformation de Lorentz	16
2	Exemples d'invariants et de quadrivecteurs	18
3	Dynamique relativiste	20
D	Normalisation des états et espace des phases	22
1	Normalisation discrète des états	23
2	Normalisation continue des états	24
3	Normalisation relativiste des états	25
E	Processus de désintégration	26
1	Règle d'or de Fermi	26
2	Forme générale du taux de désintégration	27
3	Loi exponentielle de désintégration	29
4	Relation d'incertitude temps-énergie	30
5	Annexe : distribution de Poisson	32
F	Collisions et section efficace	34
1	Notion de section efficace	34
2	Diffusion par un potentiel	35
3	Forme relativiste générale de la section efficace	39
4	Résonances et masse invariante	40
G	Problèmes	42
2	Champs quantiques et interactions	49
A	Théorie du champ scalaire	49
1	Un modèle simple en dimension 1	49
2	Limite continue	51
3	Équation de Klein-Gordon	54
4	Invariance de Lorentz de la théorie du champ scalaire	56
B	Quantification du champ scalaire	57
1	Hamiltonien	58
2	Opérateurs de création et d'annihilation	60
3	Particules en tant que quanta du champ	61
C	Interactions et diagrammes de Feynman	62
1	Hamiltonien d'interaction	62
2	Diffusion entre particules : exemple	63
3	Calcul de la section efficace	67
4	Règles de Feynman	68

5	Potentiel de Yukawa.	69
D	Problèmes.	73
3	Particules de spin 1/2	77
A	Spin et transformations de Lorentz	77
1	Rappel sur les rotations	77
2	Spineurs	79
3	Spin et invariance de Lorentz	83
4	Spineurs droits et gauches : équations de Weyl	85
5	Hélicité	87
B	Équation de Dirac	88
1	Propriétés des matrices de Dirac	90
2	Ondes planes	90
C	Quantification du champ de Dirac	93
1	Hamiltonien et relations de commutation	93
2	Fermions et anticommutateurs	95
3	Hamiltonien de Dirac	96
4	Antiparticules	97
D	Problèmes.	100
4	L'électrodynamique quantique	103
A	Le champ électromagnétique	103
1	Particule chargée dans un champ électromagnétique	103
2	Action du champ électromagnétique et équations de Maxwell.	104
3	Invariance de jauge en mécanique quantique	106
4	Couplage minimal	107
B	Photons	108
1	Quantification du champ électromagnétique	108
2	Photons	111
C	L'électrodynamique quantique.	111
1	Hamiltonien d'interaction électron-photon.	112
2	Règles de Feynman	113
3	Exemples de processus	117
D	Corrections radiatives et renormalisation	120
1	Corrections d'ordre supérieur	120
2	Constante de couplage variable	123
3	Interprétation de la renormalisation	124
E	Couplage de Yukawa et électrodynamique scalaire	126
1	Interaction de Yukawa	126
2	Champ scalaire chargé.	127
3	Théorie de Yukawa-Stückelberg de l'interaction forte	128
F	Problèmes.	130
5	Symétries	137
A	C, P et T.	138
1	La parité, ou inversion de l'espace (P)	138
2	La conjugaison de charge (C)	143

3	L'inversion du temps (T)	146
B	Isospin	150
1	Le nucléon	150
2	Représentations de l'isospin	151
3	Théorème de Noether	153
C	SU(3) et organisation des hadrons.	156
1	Étrangeté.	156
2	SU(3)	157
3	Produits tensoriels de SU(3).	160
4	Symétrie de saveur des quarks	162
D	Couleur et chromodynamique quantique	165
1	Symétrie de couleur	165
2	Théorie de Yang-Mills	169
3	Champ de jauge et générateurs	170
4	Le tenseur de Faraday	171
5	Le lagrangien de la QCD	172
6	Propriétés de la QCD	173
E	Problèmes.	177
6	Modèle standard de l'interaction électrofaible	181
A	Interaction faible	181
1	Émission bêta	181
2	Première théorie de l'interaction faible	182
3	Brisure de la parité par l'interaction faible	183
4	Théorie V-A.	185
B	Théorie de jauge SU(2) × U(1)	186
C	Mécanisme de Higgs.	188
1	Champ de Higgs	188
2	Brisure spontanée de la symétrie	189
3	Courants chargé, neutre et électromagnétique.	194
4	Relation avec la théorie V-A.	197
D	Génération des masses des fermions.	198
1	Couplage de Yukawa	198
2	Réplication des familles et mélange des quarks	200
3	Cas des leptons et des courants neutres	203
E	Physique des neutrinos.	205
1	Détection des neutrinos	205
2	Faisceaux de neutrinos	206
3	Oscillations de neutrinos	206
4	Désintégration du pion chargé.	207
F	La symétrie CP et sa brisure	209
1	Brisure de CP	211
G	L'unification des forces.	213
H	Problèmes.	217
7	Annexes	219
A	Constantes physiques et facteurs de conversion	219

B	Unités électromagnétiques	219
C	Vecteurs et tenseurs	221
D	Groupes définis par des matrices	224
E	Tableau des principales particules subatomiques	227

TABLE DES PROBLÈMES

1.1	Unités naturelles et analyse dimensionnelle	42
1.2	Désintégration du pion	42
1.3	Énergie de seuil d'une réaction	42
1.4	Cinématique des collisions	43
1.5	Variables de Mandelstam	43
1.6	Section différentielle dans le référentiel du centre d'impulsion	44
1.7	Section différentielle dans le référentiel du laboratoire	44
1.8	Diffusion d'un neutron de basse énergie	45
1.9	Droites de Kurie	45
1.10	Espace des phases	46
2.1	Équation de Proca	73
2.2	Fluctuations d'un champ scalaire	73
2.3	Oscillations entre deux types de bosons	74
2.4	Diagrammes de Feynman à l'ordre 4 pour le champ scalaire	75
2.5	Théorie en ϕ^4	75
2.6	Interaction entre deux types de bosons	75
2.7	production de particules par une source externe	76
3.1	Tenseur antisymétrique	100
3.2	Démonstration différente de l'équation de Dirac	100
3.3	Valeur moyenne du spin d'un électron	101
3.4	Application de la covariance de l'équation de Dirac	101
3.5	Paradoxe de Klein	101
4.1	Équation de Pauli	130
4.2	Champ scalaire complexe	131
4.3	Quantités physiques et opérateurs de nombre	131
4.4	Diffusion électron-muon	132
4.5	Corrections radiatives à l'effet Compton	133
4.6	Diffusion photon-photon	133
4.7	couplage de Yukawa	133
4.8	couplage de Yukawa (avec règles de Feynman)	135
5.1	Parité	177
5.2	Matrice γ^5	177
5.3	Spineurs chiraux	177
5.4	Pseudo-vecteurs et pseudo-scalaires	177
5.5	Spineur de Majorana	178
5.6	Constantes de structure	178
5.7	État neutre dans le cas de deux couleurs	178
5.8	Dérivée covariante du tenseur de Faraday	179
5.9	Champs de couleur	179
5.10	Oscillateur 3D et symétrie SU(3)	179
5.11	Produits tensoriels de SU(3)	180

TABLE DES MATIÈRES

6.1	Théorie de Fermi de l'interaction faible	217
6.2	Particule de Higgs	217
6.3	Processus de désintégration	218
6.4	Désintégration des kaons	218
6.5	Désintégration du pion	218

CHAPITRE 1

INTRODUCTION

La *physique subatomique*, comme son nom l'indique, étudie la matière organisée à des échelles de grandeurs plus petites que l'atome. En principe, cela comprend à la fois l'étude des noyaux atomiques (la *physique nucléaire*) et celle des particules «élémentaires» (la *physique subatomique* proprement dite). On préfère le terme *physique subatomique* à *physique des particules élémentaires* car beaucoup de particules subatomiques (celles qu'on appelle les *hadrons*) ne sont pas élémentaires, mais composées d'éléments plus simples (les quarks). Les deux sujets furent indissociables jusqu'aux années 1940 ; la physique nucléaire a ensuite atteint un stade de maturité relatif et s'est différenciée de la quête des constituants ultimes de la matière par ses problèmes, ses méthodes, voire ses instruments. Dans ce manuel, nous allons nous concentrer sur la physique subatomique proprement dite. Son point culminant est la description du *modèle standard*, au chapitre 6. Après ce chapitre, nous effectuons un retour dans le temps afin de décrire rapidement les concepts de base de la physique nucléaire, dans trois chapitres qui ne font pas partie du cours tel quel.

A Panorama historique

L'électron La particule connue aujourd'hui sous le nom d'*électron* est la première particule subatomique à avoir été identifiée. On fabrique, à partir des années 1860, des tubes de verre évacués dans lesquels sont installées des électrodes. Ces *tubes à décharge* sont mis au point par Heinrich GEISSLER ; le tube à décharge est ensuite amélioré par William CROOKES . On observe dans ces tubes d'étranges rayons, originaires de la cathode. Ceux-ci sont observés successivement par les savants allemands PLÜCKER, HITTORF et GOLDSTEIN. C'est ce dernier qui les appelle *rayons cathodiques* en 1876. Deux interprétations ont cours : soit ces rayons sont des particules, soit ce sont des ondes se propageant dans l'éther. Heinrich HERTZ et son étudiant Philipp LENARD croient en cette dernière hypothèse, mais Jean PERRIN soutient la première et démontre en 1895 que les rayons cathodiques portent une charge électrique négative : ils sont déviés par un champ magnétique. En 1897, Joseph John THOMSON confirme les résultats de Perrin, mesure même le rapport charge/masse de ces particules et réussit à les dévier à l'aide d'un champ électrique.¹ Plus tard, Thomson

1. Ceci n'avait pu être fait auparavant parce que le vide à l'intérieur des tubes n'était pas assez grand : l'air résiduel s'ionisait et annulait le champ électrique appliqué.

parvient à mesurer la charge de ces particules et en conclut que leur masse est environ 1 000 fois inférieure à celle de l'hydrogène (cette mesure sera raffinée par l'Américain Robert MILLIKAN en 1909, pour donner un rapport de masse plus proche de 2 000). Thomson montre aussi que tous les rayons cathodiques sont les mêmes, qu'importe la source, et donc que ces particules sont universelles. En particulier, la forme de radioactivité appelée *bêta* (β) est composée des mêmes particules. Le terme *électron* avait auparavant été suggéré par le physicien irlandais STONEY, qui avait émis l'hypothèse d'un « atome » d'électricité, en relation avec une théorie électromagnétique de l'éther. Ce terme sera rapidement adopté pour désigner les particules cathodiques de Perrin et Thomson, quoique ce dernier se refusera longtemps à l'adopter. On utilise le symbole e pour désigner l'électron.

Le photon La dualité onde-corpuscule, si fondamentale à la physique quantique, a été révélée pour la première fois en étudiant le comportement de la lumière. C'est Albert EINSTEIN qui a proposé, en 1905, que la lumière se propageait par *paquets* (*lichtquanta*), afin d'expliquer l'effet photoélectrique. Le mot *photon* est apparu plus tard (vers 1926) pour désigner ces paquets ou *quanta* et a été universellement utilisé depuis. Dans le contexte de la physique subatomique, ce sont les photons de très haute énergie (les rayons gamma) qui sont pertinents. Ceux-ci furent observés pour la première fois par Paul VILLARD en 1900. Pour cette raison, le symbole utilisé pour désigner le photon est γ .

Sur le plan théorique, le photon est une excitation du champ électromagnétique, dont la quantification a aussi été réalisée en 1926. La notion de particule et de quanta prend tout son sens lorsqu'on applique la mécanique quantique aux champs (voir chapitre 2). Cette procédure, appliquée au champ électromagnétique, mène à l'*électrodynamique quantique* (voir chapitre 4). Cette théorie cruciale, formulée pour la première fois en 1926, a dû attendre la fin des années 1940 avant de trouver un sens plus précis et a plus tard servi de cadre pour décrire toutes les interactions fondamentales.

Le proton C'est en 1919 que la particule connue sous le nom de *proton* a été formellement identifiée, par Ernest RUTHERFORD. Il a fallu pour cela se convaincre que les masses des différents noyaux étaient approximativement des multiples entiers de la masse du noyau d'hydrogène ${}^1_1\text{H}$ et qu'un objet de cette masse était transférable d'un noyau à l'autre, lors d'une *réaction nucléaire* :



Le mot *proton* (Grec pour 'premier') fut choisi pour désigner le noyau ${}^1_1\text{H}$, car il constitue selon cette hypothèse la matière première de tous les éléments.

L'antimatière En 1928, Paul DIRAC proposa une équation d'onde analogue à celle de Schrödinger, mais compatible avec la théorie de la relativité restreinte, qui décrivait l'électron et son spin. Or cette équation comporte des solutions à énergies négatives comme à énergies positives et ces solutions furent correctement interprétées par Dirac comme décrivant des particules de même masse que l'électron, mais de charge opposée. Initialement, Dirac aurait souhaité que ces particules soient les protons, ce qui lui aurait permis de décrire toutes les particules connues alors à l'aide d'une seule équation, mais il s'est rapidement convaincu que la condition d'égalité de la masse de cette particule avec celle de l'électron ne pouvait pas être levée même en considérant l'effet des interactions électromagnétiques. Il s'est alors résolu à prédire qu'une particule nouvelle,

ayant la même masse que l'électron, mais une charge positive, devait exister. Cette particule s'annihile avec un électron pour produire 2 photons, ce qui explique sa rareté naturelle. Elle fut observée pour la première fois par Carl ANDERSON dans les rayons cosmiques en 1932 et fut appelée *positron* en anglais. En français, l'usage recommandé est *positon*, mais en pratique le terme *positron* est utilisé. Ces concepts seront développés en détail au chapitre 3. Le concept d'antiparticule se généralise à toutes les particules chargées.

Le neutron Le premier modèle du noyau supposait qu'un nucléide de charge Ze et de nombre de masse A comportait A protons et $N = A - Z$ électrons. Ce modèle naïf pouvait expliquer pourquoi certains nucléides émettent des électrons (rayonnement bêta). Suite à la formulation de la mécanique quantique en 1925/1926 et en particulier du principe d'incertitude de Heisenberg, ce modèle est devenu intenable, car il était clair qu'on ne pouvait confiner des électrons dans un espace aussi restreint que le noyau atomique ($\sim 10^{-15}\text{m}$) sans entraîner des fluctuations d'énergie énormes pour ces électrons, fluctuations qui auraient immédiatement éjecté les électrons du noyau. Cette situation gênante perdura jusqu'en 1932, quand un nouveau type de «rayon» très pénétrant et capable d'éjecter des protons par collision fut découvert. Ce rayon fut correctement interprété par James CHADWICK comme constituant une particule neutre d'une masse légèrement supérieure à celle du proton. Cette découverte a permis l'élaboration d'une véritable physique nucléaire. Une nouvelle force fondamentale, simplement appelée l'*interaction forte* ou *force nucléaire forte*, doit être invoquée pour expliquer la cohésion du noyau, c'est-à-dire l'attraction des protons et des neutrons entre eux.

Les neutrinos La désintégration bêta, l'une des trois formes de radioactivité découvertes à l'aube du XX^e siècle, a cette propriété particulière qu'elle ne semblait pas respecter les lois de conservations chères à la physique (énergie, quantité de mouvement et moment cinétique). Dans sa forme la plus courante, la désintégration bêta prend apparemment la forme suivante : $n \rightarrow p + e$ (le neutron se désintègre en un proton et un électron). Or, dans tout processus où une seule particule se désintègre en deux fragments, la conservation de l'énergie et de la quantité de mouvement impose que chacun des deux fragments ait une énergie bien définie dans le référentiel du neutron. Ce n'est pas le cas cependant : on observe un spectre continu d'énergie pour l'électron émis (voir Fig. 6.1 à la page 182). La solution à ce paradoxe, proposée par Wolfgang PAULI, a été de supposer qu'une troisième particule, électriquement neutre et très peu massive, faisait partie des produits de la désintégration bêta. Cette particule fut appelée *neutrino* par Enrico FERMI. Il s'agit du cas classique d'une particule hypothétique et pratiquement inobservable, introduite dans le but de respecter des lois de conservation fondamentales. Ce n'est qu'en 1955 que le neutrino put être observé via sa collision avec d'autres particules.

Plus important, le processus de désintégration bêta n'est régi ni par l'interaction électromagnétique, ni par l'interaction forte, mais par une nouvelle interaction, appelée *interaction faible* ou *force nucléaire faible*, décrite pour la première fois par Fermi en 1933. Une description plus correcte de l'interaction faible, élaborée dans les années 1960, formera la base du modèle standard. Elle est exposée au chapitre 6. En fait, il s'agit d'une synthèse de l'interaction électromagnétique et de l'interaction faible, décrite par la *théorie électrofaible*, attribuée à Sheldon GLASHOW (1961), Steven WEINBERG et Abdus SALAM (1967).

Les mésons Avec l'électrodynamique quantique est apparu le concept de *particule médiatrice* : chaque force fondamentale est représentée par une particule qui est, en un certain sens expliqué au chapitre 2, échangée entre les particules qui sont sous l'influence de cette interaction. Dans cette représentation, des particules chargées en interaction échangent entre elles des photons *virtuels*, une manifestation des états virtuels en théorie quantique des perturbations. L'idée d'associer une particule nouvelle à l'interaction nucléaire forte a été proposée par Hideki YUKAWA et Ernst STÜCKELBERG en 1934 et 1935. Cette particule hypothétique, d'une masse située entre 100 MeV et 200 MeV, fut appelée *méson*. La découverte de particules ayant à peu près ces caractéristiques dans les rayons cosmiques en 1936, par Carl Anderson et Seth NEDDERMEYER, fut suivie d'une période de confusion causée par le fait que ladite particule ne semblait pas interagir fortement. On s'aperçut plus tard que la particule découverte par Anderson et Neddermeyer n'avait rien à voir avec l'interaction forte, mais était plutôt une espèce d'électron lourd qu'on appela *méson μ* ou *muon*. Une particule plus proche de celle envisagée par Yukawa fut découverte en 1947 et reçut le nom de *méson π* ou *pion*.

Les quarks La mise en service d'accélérateurs de particules plus puissants dans les années 1950 multiplia le nombre de nouvelles particules sensibles aux interactions fortes. Ces particules, collectivement appelées *hadrons* étaient si nombreuses qu'il semblât inévitable qu'elles soient non pas élémentaires, mais au contraire composées d'éléments plus simples. En 1963, Murray GELL-MANN proposa que ces hadrons soient composés de trois variétés de particules élémentaires qu'il appelle *quarks*; il s'agissait à cette époque des quarks *u*, *d* et *s*. Plus tard on ajouta à cette liste les quarks *c*, *b* et *t*. Gell-Mann donna à sa théorie, fondée uniquement sur des considérations de symétrie, le nom de *eightfold way* (ou *voie octuple*). Elle sera décrite à la section 5.C. Bien que des évidences expérimentales tendent à confirmer le caractère composite des hadrons dans les années 1960, c'est la découverte d'une nouvelle série de mésons en novembre 1974 qui convainquit la communauté scientifique de l'existence des quarks. Parallèlement à cette découverte, une théorie fondamentale de l'interaction entre les quarks, la *chromodynamique quantique* ou QCD, sorte de généralisation de l'électrodynamique quantique, fut proposée au début des années 1970 par David GROSS, Frank WILCEK et David POLITZER. Cette théorie est décrite à la section 5.D.

Le boson de Higgs Les arguments basés sur la symétrie sont d'une importance capitale dans le modèle standard. S'ajoutent à cela des arguments basés sur le *principe de naturalité*, qui stipule que les paramètres d'une théorie ne devraient pas être ajustés finement afin de produire des coïncidences artificielles. Le modèle standard, pour être *naturel*, doit comporter au départ des particules qui, pour des raisons de symétrie, n'ont initialement pas de masse. Leur masse doit être générée suite à une *brisure spontanée* de cette symétrie, constituant le *mécanisme de Higgs-Anderson*, qui sera expliqué à la section 6.C de ce manuel. Cette théorie fut proposée à peu près en même temps par plusieurs théoriciens en 1964, mais le nom de Peter Higgs est celui qui lui est resté collé. Elle prédit l'existence d'une particule très massive, le *boson de Higgs*, sorte de sous-produit de ce mécanisme de génération des masses, mais qui autrement joue un rôle relativement effacé dans le modèle standard. La quête de cette particule a été un enjeu majeur de la physique subatomique expérimentale dans années 2000 et a justifié pratiquement à elle seule la construction du *Large Hadron Collider*. Elle a finalement été observée en 2012. Il s'agit de la dernière découverte majeure en physique subatomique.

B Le modèle standard : survol

Toute explication scientifique, simple ou complexe, repose sur ce qu'on appelle un *paradigme*, c'est-à-dire un modèle avec ses hypothèses et concepts fondamentaux. Une « super-théorie », en quelque sorte. La physique subatomique repose sur un tel paradigme, couramment appelé le *modèle standard*².

Ce modèle standard est provisoirement considéré comme la théorie correcte des phénomènes subatomiques. « Provisoirement » signifie depuis le milieu des années 1970, et jusqu'à ce qu'une théorie plus satisfaisante soit proposée, ce qui n'est toujours pas le cas au moment d'écrire ce texte.

À la base du modèle standard est un outil théorique fondamental : la *théorie quantique des champs* (TQC). La TQC est une description quantique d'objets qui s'étendent dans tout l'espace, comme le champ électromagnétique, qui en est l'exemple le plus ancien. La TQC doit aussi être formulée dans le cadre de la relativité restreinte, car toute description des particules élémentaires se doit de respecter le principe de relativité.

Le modèle standard offre une description des phénomènes subatomiques compatible avec essentiellement toutes les données expérimentales, exception faite des oscillations de neutrinos soupçonnées depuis les années 1980 et confirmées plus solidement au début des années 2000. Il décrit les interactions forte, faible et électromagnétique à l'aide de ce qu'on appelle des *théories de jauge*. Ces interactions agissent sur des particules de matière, des fermions, qui sont rangées en trois familles semblables et comptent quarks et leptons (voir figure 1.1).

1.B.1 Classification des particules élémentaires

Le concept fondamental de la théorie quantique des champs est que les particules élémentaires sont considérées comme des *quanta*, c'est-à-dire des oscillations ou fluctuations quantifiées de *champs*. Ces particules (et par extension, les champs correspondants) sont soit des *bosons* ou des *fermions*. Les champs de bosons, tel le champ électromagnétique, peuvent être formulés plus intuitivement dans la limite classique ; ils correspondent aux interactions fondamentales. Les champs de fermions n'ont pas de limite classique ; ce sont des objets purement quantiques qui correspondent aux particules de matière (quarks et leptons).

2. ou, plus longuement, *modèle standard des particules élémentaires*, pour le distinguer d'autres modèles standards qui prévalent dans d'autres disciplines comme la cosmologie ou la biologie moléculaire...

			Q	famille 1	famille 2	famille 3
bosons de jauge	γ	photon (int. électromagnétique)	0	ν_e (neutrino e)	ν_μ (neutrino μ)	ν_τ (neutrino τ)
	W^\pm, Z	(int. faible)	-1	e (électron)	μ (muon)	τ (taupon)
	g	gluons (int. forte)	$\frac{2}{3}$	u (up)	c (charm)	t (top)
	H	Higgs	$-\frac{1}{3}$	d (down)	s (strange)	b (bottom)
bosons (forces)				fermions (matière)		

FIGURE 1.1

Tableau des particules élémentaires dans le modèle standard. Les bosons figurent à gauche, les fermions à droite. La charge électrique Q des différents fermions est indiquée. À chaque espèce de fermion est aussi associée une antiparticule.

Fermions Commençons par les fermions, tous de spin $\frac{1}{2}$ comme l'électron. Ils sont organisés en trois familles distinctes, semblables dans la manière avec laquelle ils interagissent (par exemple la charge électrique Q), mais très différentes dans leurs masses (les familles sont numérotées dans l'ordre croissant des masses). On distingue les *leptons*,³ qui ne ressentent pas l'interaction forte, des *quarks*, qui la ressentent et qui, par une propriété singulière de l'interaction forte appelée *confinement*, ne peuvent être observés en tant que particules libres. Chaque quark, de surcroît, existe en trois *couleurs*, c'est-à-dire trois exemplaires équivalents (non répétés sur le tableau). En tenant compte de ce fait, on voit que la charge électrique totale de chaque famille est nulle.

Les fermions élémentaires ont tous un spin $\frac{1}{2}$ et sont tous décrits par l'équation de Dirac (c'est-à-dire que les champs correspondants sont des *champs de Dirac*, mathématiquement décrits par des spineurs à 4 composantes). Ils ont tous des antiparticules (non inscrites sur le tableau), qu'on note généralement par une barre au-dessus du symbole (par exemple, \bar{e} désigne l'antiélectron, ou positron).

Les quarks, qui ne sont pas observables individuellement, forment des particules composites de deux types : les *mésons* sont formés d'un quark et d'un antiquark (on écrirait $q\bar{q}$, q signifiant un quark quelconque) et les *baryons*, formés de trois quarks (qqq). Les mésons sont des bosons, car ils sont formés de deux fermions, mais ne sont pas élémentaires. Les baryons, dont les plus connus sont le proton (uud) et le neutron (udd), sont des fermions. Les baryons et les mésons sont collectivement appelés *hadrons*,⁴ terme qui désigne toutes les particules sensibles à l'interaction forte.

La nature fermionique des quarks et des leptons, c'est-à-dire le fait qu'ils obéissent au principe de Pauli, donne justement à la matière la propriété de solidité, d'impénétrabilité qui la distingue des ondes électromagnétiques et, par extension, des objets formés de bosons élémentaires seulement.

3. Du grec *leptos*, qui veut dire 'maigre', 'mince' ou 'délicat'.

4. Du grec *adros*, qui veut dire 'grand', 'fort'.

Bosons Les bosons sont associés aux interactions. Ceci signifie que les fermions n'interagissent pas directement entre eux, mais uniquement par l'intermédiaire des bosons. Un concept important qui survient dans la théorie des perturbations appliquée à la TQC est celui de *particule virtuelle*. Une particule virtuelle n'est pas observable et ne constitue pas un objet se propageant librement, mais constitue plutôt un objet transitoire, une fluctuation du champ correspondant, dont l'existence se définit sur un temps très court. On dit que les fermions interagissent en échangeant (c'est-à-dire en émettant et absorbant) des bosons virtuels.

Les trois interactions fondamentales décrites dans le modèle standard (celui-ci ignore la gravité) sont :

1. L'électromagnétisme. La particule correspondante est le *photon* (γ). Ceci signifie que les oscillations quantifiées du champ électromagnétique forment des objets détectables qu'on appelle photons et que les particules chargées peuvent échanger des photons virtuels.
2. L'interaction faible. Les particules correspondantes n'ont pas de nom, mais sont notées W^\pm et Z . La faiblesse de l'interaction faible aux basses énergies est liée à la masse très grande de ces particules (de l'ordre de 100 GeV). En fait, l'interaction faible et l'interaction électromagnétique sont intimement reliées et forment ensemble l'interaction dite *électrofaible*.
3. L'interaction forte. La particule correspondante est le *gluon* (g). Ceux-ci (car il y en a huit) sont inobservables à l'état isolé, comme les quarks, car ils sont soumis à la règle du confinement. En principe, des objets composés de deux gluons ou de trois gluons (les boules de glue ou *glueballs*) ne sont pas interdits par cette règle, mais aucun n'a été observé. Cependant, l'existence des gluons laisse des traces dans les expériences de diffusion de protons à haute énergie.

Ces trois interactions reposent sur un principe théorique important appelé *invariance de jauge*. En somme, les théories de ces interactions sont modelées sur l'interaction électromagnétique, quoique la théorie électrofaible et la théorie des interactions fortes (appelée *chromodynamique quantique* ou QCD) soient plus complexes.

L'un des ingrédients de la théorie électrofaible est le *mécanisme de Higgs* qui explique l'origine de la masse des fermions et celle des W^\pm et Z . Ce mécanisme entraîne l'existence d'un boson supplémentaire, dit *boson de Higgs* (noté H), qui vient probablement d'être observé dans deux expériences du LHC (CERN). Le boson de Higgs a un spin nul, alors que les bosons de jauge ont nécessairement un spin unité.

De tous ces bosons et fermions élémentaires, seuls l'électron, le photon et les trois espèces de neutrinos existent de manière stable en tant que particules individuelles. Tous les autres sont instables ou n'existent qu'à l'intérieur d'objets composites. Le proton est le seul baryon stable. Aucun méson stable n'existe. Bien sûr, des neutrons existent à l'intérieur des noyaux, car leur environnement immédiat change les conditions de stabilité. De même, de la matière « étrange » (c'est-à-dire comportant des particules composées en partie de quarks s) existe peut-être dans le coeur des étoiles à neutrons, là encore parce que les canaux de désintégration de ces objets sont « obstrués » par la préexistence des produits de désintégration dans l'environnement.

C Rappels de relativité restreinte

La physique des hautes énergies repose en partie sur la théorie de la relativité restreinte. Le but de cette section est de procéder à des rappels de base sur cette théorie. La notion de quadrivecteur et la notation associée sont particulièrement importantes.

Avertissement : unités naturelles

À partir de ce point, nous utiliserons le système d'unités naturelles, dans lequel $c = 1$ et $\hbar = 1$. Dans ce système d'unités, une seule dimension fondamentale indépendante demeure : une longueur est équivalente à un temps, et à une énergie inverse ou une masse inverse. Il est toujours possible, dans une expression analytique, de restaurer c et \hbar par analyse dimensionnelle. En particulier, on peut utiliser le facteur de conversion $\hbar c = 197 \text{ MeVfm}$ pour retrouver des valeurs numériques. Dans ce système d'unités, la constante de structure fine a l'expression $\alpha = e^2/4\pi \approx 1/137$.

1.C.1 Transformation de Lorentz

L'espace euclidien tridimensionnel, auquel on ajoute la dimension temporelle, forme ce qu'on appelle l'*espace-temps*. Les points dans l'espace-temps sont appelés *événements*. On peut introduire dans l'espace-temps une base notée e_μ ($\mu = 0, 1, 2, 3$) où le vecteur e_0 est un vecteur unitaire dans la direction temporelle. Un changement de référentiel est en fait une transformation qui nous amène vers une nouvelle base e'_μ . Il existe un produit scalaire bien défini sur l'espace-temps, défini alors par le tenseur métrique suivant :

$$g_{\mu\nu} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix} \quad (1.2)$$

(voir l'annexe 7.C pour une revue des vecteurs et tenseurs, ainsi que des indices covariants et contravariants). En fonction des coordonnées $(t, x, y, z) = (x^0, x^1, x^2, x^3)$, qui sont les composantes contravariantes du quadrivecteur position de l'événement $x = x^\mu e_\mu$, la norme de x s'exprime comme

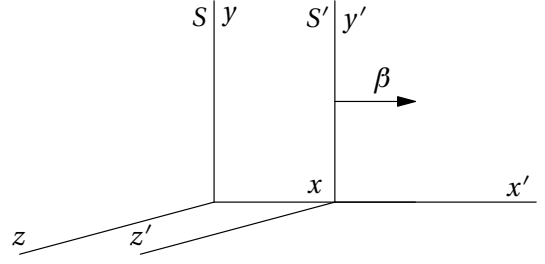
$$x^2 = (x^0)^2 - (x^1)^2 - (x^2)^2 - (x^3)^2 \quad (1.3)$$

et porte le nom d'*intervalle* (sous-entendu : intervalle entre l'événement x et l'origine).

Considérons deux référentiels S et S' , dont les axes cartésiens sont parallèles, se déplaçant l'un par rapport à l'autre à une vitesse v le long de l'axe des x et dont les origines coïncident au temps $t = 0$

FIGURE 1.2

Disposition des axes cartésiens pertinents à la transformation de Lorentz (1.4)



(Fig. 1.2). Les coordonnées et le temps des deux référentiels sont reliés par la *transformation de Lorentz* :

$$\begin{aligned} x' &= \gamma(x - vt) & y' &= y \\ t' &= \gamma(t - vx) & z' &= z \end{aligned} \quad (1.4)$$

où $\gamma = 1/\sqrt{1-v^2}$. On introduit aussi la *rapidité* η définie par la relation

$$\tanh \eta = \beta \quad (1.5)$$

La transformation de Lorentz prend alors la forme

$$\begin{aligned} x' &= x \cosh \eta - t \sinh \eta & y' &= y \\ t' &= t \cosh \eta - x \sinh \eta & z' &= z \end{aligned} \quad (1.6)$$

L'avantage de la rapidité est que (i) la transformation de Lorentz est formellement similaire à une rotation dans l'espace, où les fonctions circulaires ont été remplacées par des fonctions hyperboliques et (ii) la composition de deux transformations de Lorentz successives dans la même direction se fait par simple addition des rapidités.

La transformation de Lorentz (1.4) permet de relier entre elles les composantes décrivant le même événement dans deux référentiels différents. Lors d'une transformation de Lorentz générale, la coordonnée spatio-temporelle x^μ se transforme comme suit :

$$x'^\mu = \Lambda^\mu{}_\nu x^\nu \quad (1.7)$$

où $\Lambda^\mu{}_\nu$ est une matrice de transformation. L'indice de gauche numérote les rangées, l'indice de droite les colonnes et les indices répétés sont sommés. Par exemple, pour la transformation (1.4) la matrice de transformation est

$$= \begin{pmatrix} \gamma & -v\gamma & 0 & 0 \\ -v\gamma & \gamma & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} = \begin{pmatrix} \cosh \eta & -\sinh \eta & 0 & 0 \\ -\sinh \eta & \cosh \eta & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \quad (1.8)$$

mais elle serait différente pour un changement de référentiel accompagné d'une rotation des axes, ou suivant un axe quelconque. La matrice est l'équivalent de la matrice \tilde{S} décrite à l'annexe 7.C.

La relation (1.4) laisse invariante l'expression de l'*intervalle* associé à la quadri-position x :

$$(x^0)^2 - (x^1)^2 - (x^2)^2 - (x^3)^2 = (x'^0)^2 - (x'^1)^2 - (x'^2)^2 - (x'^3)^2 \quad (1.9)$$

Une matrice appartiendra, par définition, au *groupe de Lorentz* si elle préserve cette forme de l'intervalle. Cela revient à dire que les composantes du tenseur métrique (1.2) (ou de son inverse $g^{\mu\nu}$) doivent être les mêmes dans tous les référentiels. Autrement dit,

$$g^{\mu\nu} = \Lambda^\mu{}_\alpha \Lambda^\nu{}_\beta g^{\alpha\beta} \quad (1.10)$$

Cette condition doit être respectée par une matrice pour qu'elle appartienne au groupe de Lorentz. L'équation ci-haut représente 10 conditions indépendantes, car elle est symétrique en (μ, ν) . Comme la matrice comporte a priori 16 composantes indépendantes, cela laisse six paramètres libres qui peuvent spécifier un élément du groupe de Lorentz. Ces six paramètres correspondent aux trois composantes de la vitesse relative entre deux référentiels, ainsi qu'à trois paramètres de rotation entre les axes cartésiens des deux référentiels (trois angles d'Euler, par exemple).

Selon la relation (7.11), les composantes covariantes se transforment à l'aide de la transposée de la matrice inverse de , soit

$$A'_\mu = (\Lambda^{-1})^\nu{}_\mu A_\nu \quad (1.11)$$

La matrice inverse de (1.8) s'obtient simplement en changeant le signe de v .

1.C.2 Exemples d'invariants et de quadrivecteurs

Quadrigradient Examinons maintenant comment se transforme l'opérateur gradient augmenté de la dérivée par rapport au temps, noté

$$\partial_\mu = \frac{\partial}{\partial x^\mu} = \left(\frac{\partial}{\partial t}, \nabla \right) \quad (1.12)$$

Sachant que $x^\nu = (\Lambda^{-1})^\nu{}_\mu x'^\mu$ (transformation inverse), on voit que

$$\partial'_\mu = \frac{\partial x^\nu}{\partial x'^\mu} \frac{\partial}{\partial x^\nu} = (\Lambda^{-1})^\nu{}_\mu \partial_\nu \quad (1.13)$$

et donc que le quadrigradient se transforme comme les composantes covariantes d'un quadrivecteur, ce qui justifie la notation utilisée.

Temps propre Le *temps propre* τ d'un objet en mouvement est simplement le temps tel qu'il s'écoule dans le référentiel de l'objet. Dans le référentiel inertiel S' qui se déplace avec l'objet à un instant donné, la différentielle de temps propre coïncide avec dt' . Cependant, l'objet peut en général être accéléré de sorte que le référentiel S' n'est pas le même à tous les instants. Quel que soit le référentiel dans lequel on observe l'objet, la différentielle de temps propre peut toujours être exprimée comme suit :

$$d\tau = \sqrt{dx^\mu dx_\mu} = dt \sqrt{1 - v^2} = \frac{dt}{\gamma} \quad (1.14)$$

où v est la vitesse de l'objet, qui peut dépendre du temps. En effet, dans le référentiel S' , $dx^\mu dx_\mu = d\tau^2$ et de plus l'expression ci-haut est invariante, puisque $dx^\mu dx_\mu$ est un invariant. Pour l'obtenir le temps propre total écoulé sur la trajectoire de l'objet, on intègre $d\tau$:

$$\tau = \int d\tau = \int \frac{dt}{\gamma} = \int dt \sqrt{1 - v^2} \quad (1.15)$$

Quadrivitesse Considérons un objet se déplaçant à une vitesse instantanée \mathbf{v} . On définit sa *quadrivitesse* u^μ comme

$$u^\mu = \frac{dx^\mu}{d\tau} \quad (1.16)$$

La différentielle $d\tau$ est invariante, alors que dx^μ se transforme comme les coordonnées; donc u^μ est un quadrivecteur. En fonction de la vitesse \mathbf{v} , les composantes explicites sont

$$u^\mu : (\gamma, \gamma \mathbf{v}) \quad (1.17)$$

Par contraction on obtient l'invariant $u^\mu u_\mu = 1$. Le temps propre écoulé le long d'une trajectoire peut alors s'écrire comme

$$\tau = \int d\tau = \int u_\mu dx^\mu \quad (1.18)$$

Quadriaccélération On définit aussi la quadriaccélération

$$a^\mu = \frac{du^\mu}{d\tau} = \gamma \frac{du^\mu}{dt} \quad (1.19)$$

L'expression explicite (1.17) de la quadrivitesse donne

$$a^\mu = \left(\frac{(\mathbf{v} \cdot \mathbf{a})}{(1 - \mathbf{v}^2)^2}, \frac{\mathbf{a}}{1 - \mathbf{v}^2} + \frac{(\mathbf{v} \cdot \mathbf{a})\mathbf{v}}{(1 - \mathbf{v}^2)^2} \right) \quad (1.20)$$

À partir de ce quadrivecteur, on peut construire certains invariants. Par exemple,

$$u_\mu a^\mu = u_\mu \frac{du^\mu}{d\tau} = \frac{1}{2} \frac{d}{d\tau} (u_\mu u^\mu) = 0 \quad (1.21)$$

En utilisant les expressions explicites pour u^μ et a^μ en fonction des composantes temporelles et spatiales, la condition $u_\mu a^\mu = 0$ ressort de manière triviale. Un autre invariant à considérer est $a^\mu a_\mu$. On calcule que

$$a^\mu a_\mu = -\gamma^4 (\mathbf{a}^2 + \gamma^2 (\mathbf{v} \cdot \mathbf{a})^2) \quad (1.22)$$

Cette quantité est la même dans tous les référentiels, même si les valeurs de \mathbf{v} et de \mathbf{a} dépendent du référentiel.

Quadricourant Un autre quadrivecteur est formé par la densité de courant et la densité de charge :

$$J^\mu : (\rho, \mathbf{J}) \quad (1.23)$$

Pour s'en convaincre, il suffit d'écrire l'équation de continuité comme

$$\nabla \cdot \mathbf{J} + \frac{\partial \rho}{\partial t} = \partial_\mu J^\mu = 0 \quad (1.24)$$

Comme la conservation de la charge doit être valide dans tous les référentiels et que ∂_μ est un quadrivecteur, le quadricourant J^μ en est forcément un lui aussi. Une autre façon de se convaincre que J^μ est un quadrivecteur est de considérer l'expression de la densité de courant associée à un ensemble de charges ponctuelles en mouvement :

$$\mathbf{J}(\mathbf{r}) = \sum_i q^{(i)} \mathbf{v}^{(i)} \delta(\mathbf{r} - \mathbf{r}^{(i)}) \quad (1.25)$$

Ici $q^{(i)}$, $\mathbf{r}^{(i)}$ et $\mathbf{v}^{(i)}$ sont respectivement la charge, la position et la vitesse de la i^e particule. La fonction delta dans la formule ci-haut n'est pas un invariant de Lorentz, puisqu'elle est définie par la relation

$$\int d^3r^{(i)} \delta(\mathbf{r} - \mathbf{r}^{(i)}) = 1 \quad (1.26)$$

valide dans tous les référentiels et que d^3r n'est pas invariant. Cependant, $\gamma^{(i)} d^3r^{(i)}$ est un invariant, où $\gamma^{(i)} = 1/\sqrt{1-(v^{(i)})^2}$. La fonction $(1/\gamma^{(i)})\delta(\mathbf{r} - \mathbf{r}^{(i)})$ est donc aussi un invariant. Écrivons donc la densité de courant et la densité de charge comme

$$\mathbf{J}(\mathbf{r}) = \sum_i q^{(i)} \gamma^{(i)} \mathbf{v}^{(i)} [\delta(\mathbf{r} - \mathbf{r}^{(i)})/\gamma^{(i)}] \quad (1.27)$$

$$\rho(\mathbf{r}) = \sum_i q^{(i)} \gamma^{(i)} [\delta(\mathbf{r} - \mathbf{r}^{(i)})/\gamma^{(i)}] \quad (1.28)$$

En fonction de la quadrivitesse de chaque particule, on a

$$J^\mu = \sum_i q^{(i)} u^{(i)\mu} [\delta(\mathbf{r} - \mathbf{r}^{(i)})/\gamma^{(i)}] \quad (1.29)$$

Comme $q^{(i)}$ et la fonction delta modifiée sont des invariants, on conclut que J^μ est bel et bien un quadrivecteur.

1.C.3 Dynamique relativiste

Action d'une particule libre En relativité, les équations du mouvement doivent être les mêmes dans tous les référentiels inertiels. Pour cela il faut que l'action soit invariante par rapport aux transformations de Lorentz, c'est-à-dire qu'elle ait la même forme dans tous les référentiels. Ainsi, si elle est stationnaire dans un référentiel, elle le sera dans tous les référentiels. Guidés par ce principe, essayons d'obtenir l'action d'une particule libre. Soit $x^\mu(\tau)$ la trajectoire de la particule dans l'espace-temps, en fonction du temps propre τ . La quantité invariante la plus simple qu'on peut construire avec une telle trajectoire est sa longueur propre, c'est-à-dire le temps écoulé dans le référentiel (non inertielle, en général) de la particule. Cette quantité doit être multipliée par une constante qui donne à l'action ses unités naturelles, soit celles d'une énergie fois un temps. La seule constante relative à la particule ayant les unités d'une énergie est sa masse m . On suppose donc que l'action est

$$S_0 = -m \int d\tau = -m \int u_\mu dx^\mu \quad (1.30)$$

où u_μ est la quadrivitesse de la particule (la constante multiplicative -1 a été ajoutée afin de retrouver le résultat non relativiste). Puisque $d\tau = dt/\gamma$, on peut écrire

$$S_0 = -m \int dt \sqrt{1-v^2} \quad (1.31)$$

Le lagrangien d'une particule libre est donc

$$L_0 = -m \sqrt{1-v^2} \quad (1.32)$$

Dans l'approximation non relativiste ($v \ll 1$) on a $\sqrt{1-v^2} \approx 1 - v^2/2$ et donc

$$L_0 \approx -m + \frac{1}{2} m v^2 \quad (1.33)$$

On retrouve donc le lagrangien habituel, plus une constante ($-m$) qui n'affecte pas les équations du mouvement.

L'impulsion généralisée \mathbf{p} est définie comme

$$\mathbf{p} = \frac{\partial L_0}{\partial \mathbf{v}} = \frac{m\mathbf{v}}{\sqrt{1-v^2}} \quad (1.34)$$

La fonction de Hamilton (hamiltonien) qui correspond à ce lagrangien est

$$\begin{aligned} H_0 &= \mathbf{p} \cdot \mathbf{v} - L_0 \\ &= \frac{m\mathbf{v}^2}{\sqrt{1-v^2}} + m\sqrt{1-v^2} \\ &= \frac{m}{\sqrt{1-v^2}} \end{aligned} \quad (1.35)$$

Le hamiltonien doit en principe être exprimé en fonction de l'impulsion conjuguée \mathbf{p} , ce qui donne dans ce cas

$$H = \sqrt{\mathbf{p}^2 + m^2} \quad (1.36)$$

Le hamiltonien est dans ce cas égal à l'énergie cinétique de la particule, modulo une constante additive. Dans l'approximation non relativiste, H devient

$$H_0 \approx m + \frac{\mathbf{p}^2}{2m} \quad (1.37)$$

On reconnaît l'énergie cinétique usuelle, plus une constante : l'énergie au repos m .

Quadri-impulsion On définit le quadrivecteur impulsion d'une particule massive comme la masse multipliée par la quadrivitesse :

$$p^\mu = m u^\mu = (m\gamma, m\gamma\mathbf{v}) \quad (1.38)$$

On constate, d'après l'expression de l'énergie E et de la quantité de mouvement \mathbf{p} de la particule, que les composantes de la quadri-impulsion s'expriment également ainsi :

$$p^\mu = (E, \mathbf{p}) \quad (1.39)$$

et que le carré de ce quadrivecteur n'est autre que la masse au carré : $p_\mu p^\mu = m^2$, ou encore

$$\boxed{E^2 - \mathbf{p}^2 = m^2} \quad (1.40)$$

Lors d'une collision entre particules, la somme des quadrivecteurs énergie-impulsion de toutes les particules impliquées est conservée, c'est-à-dire qu'elle est la même après et avant la collision. Par

exemple, si deux particules (numérotées 1 et 2) entrent en collision pour produire deux autres particules (numérotées 3 et 4) et que les quadri-impulsions associées sont notées $p_{1,2,3,4}^\mu$, la conservation de l'énergie et de la quantité de mouvement s'écrit

$$p_1^\mu + p_2^\mu - p_3^\mu - p_4^\mu = 0 \quad (1.41)$$

Comme l'expression ci-dessus est un quadrivecteur dont toutes les composantes sont nulles, les dites composantes sont nulles dans tous les référentiels et la conservation de l'énergie-impulsion est alors valide indépendamment du référentiel utilisé.

La propagation d'une onde fait intervenir le quadrivecteur d'onde $k^\mu = (\omega, \mathbf{k})$, qui ne diffère de la quadri-impulsion que par un facteur \hbar ($p^\mu = \hbar k^\mu$). La distinction entre quadri-impulsion et quadrivecteur d'onde n'apparaît donc pas dans le système des unités naturelles. Une onde plane ψ se propageant avec un quadrivecteur d'onde p^μ aura donc la forme suivante (notez le signe de la phase) :

$$\psi(\mathbf{r}, t) = \psi_0 e^{-i p_\mu x^\mu} = \psi_0 e^{i(\mathbf{k} \cdot \mathbf{r} - \omega t)} \quad (1.42)$$

la relation (1.40), qui lie la fréquence au vecteur d'onde, est alors naturellement désignée sous le nom de *relation de dispersion*. La phase est alors un invariant de Lorentz, ce qui est naturel étant donné que les phénomènes d'interférence sont observés indépendamment du référentiel.

Notons enfin que les facteurs ν et γ figurant dans la transformation de Lorentz peuvent être exprimés en fonction de l'énergie et de l'impulsion d'une particule se déplaçant à la vitesse ν :

$$\boxed{\nu = \frac{|\mathbf{p}|}{E} \quad \gamma = \frac{E}{m}} \quad (1.43)$$

Particules de masse nulle La relation (1.40) reste valable dans la limite d'une masse nulle. On trouve alors $E = |\mathbf{p}|$. Physiquement, il est impossible de distinguer une particule de masse exactement nulle d'une particule dont la masse est infinitésimale. En pratique, toute particule dont l'énergie est très grande par rapport à sa masse se comporte approximativement comme si sa masse était nulle : on dit alors que la particule est *ultra-relativiste*. Le prototype de particule de masse nulle est le photon.

D Normalisation des états et espace des phases

Qu'est-ce qu'une particule? Intuitivement, il s'agit d'un objet qui se propage librement pendant un temps suffisamment long avant de se désintégrer, ou d'entrer en collision avec d'autres objets. «Se propager librement» signifie qu'en première approximation, la dynamique de l'objet est dominée par l'énergie cinétique et que les états quantiques pertinents pour décrire l'objet sont des ondes planes, c'est-à-dire des états à impulsion \mathbf{p} définie.

La normalisation de ces états est matière à convention. Trois de ces conventions seront utilisées ici : une normalisation discrète (ND), une normalisation continue (NC) et une normalisation relativiste (NR) invariante de Lorentz.

1.D.1 Normalisation discrète des états

Dans la normalisation discrète, on suppose que l'espace physique est contenu dans une boîte rectangulaire dont les côtés sont de longueurs L_x , L_y et L_z , et on impose des conditions aux limites périodiques aux fonctions d'onde. Une onde plane a alors la forme (non normalisée)

$$\psi(\mathbf{r}) = e^{i\mathbf{p}\cdot\mathbf{r}} = e^{ip_x x} e^{ip_y y} e^{ip_z z} \quad (1.44)$$

La condition de périodicité restreint les valeurs possibles de l'impulsion. Il faut en effet que la translation $x \rightarrow x + L_x$ (ou l'équivalent en y et z) n'affecte pas la fonction d'onde. Donc,

$$e^{ip_x L_x} = 1 \implies p_x = \frac{2\pi n_x}{L_x} \quad (1.45)$$

où n_x est un entier. Les impulsions possibles sont donc

$$\mathbf{p} = 2\pi \left(\frac{n_x}{L_x}, \frac{n_y}{L_y}, \frac{n_z}{L_z} \right) \quad (1.46)$$

où n_x , n_y et n_z peuvent prendre toutes les valeurs entières possibles ($\in \mathbb{Z}$). Une seule valeur de p_x est admise dans un intervalle $\Delta p_x = 2\pi/L_x$, correspondant à $\Delta n_x = 1$. Il y a donc un seul état (une seule valeur possible des entiers) dans un élément de volume donné par

$$\Delta p_x \Delta p_y \Delta p_z = \frac{(2\pi)^3}{\mathcal{V}} \quad (1.47)$$

où $\mathcal{V} = L_x L_y L_z$ est le volume de l'espace physique. Le nombre d'états dans un élément de volume $d^3 p$ est donc simplement le rapport

$$\frac{\mathcal{V}}{(2\pi)^3} d^3 p \quad (1.48)$$

Exprimé autrement, le nombre d'états d'onde plane par élément de volume $d^3 p$ de l'espace des impulsions est

$$\frac{d^3 n}{d^3 p} = \frac{\mathcal{V}}{(2\pi)^3} \quad (1.49)$$

Notons qu'on peut aussi exprimer ce résultat comme un nombre d'états par unité de volume dans l'espace des phases, c'est-à-dire l'espace décrit par les positions (\mathbf{r}) et les impulsions (\mathbf{p}) :

$$\frac{d^3 n}{d^3 p d^3 r} = \frac{1}{(2\pi)^3} \quad \text{ou} \quad = \frac{1}{h^3} \quad (1.50)$$

si on retourne aux unités usuelles. C'est sous cette forme que le résultat est le plus simple à mémoriser : une cellule d'espace des phases de volume h^3 (ou $(2\pi)^{-3}$ en unités naturelles) contient exactement un état quantique.

Les états d'impulsion donnée, dans cette normalisation, forment un ensemble discret : les valeurs de \mathbf{p} sont nettement séparées. Les états sont donc normalisables et on peut imposer la condition

$$\langle \mathbf{p}' | \mathbf{p} \rangle = \delta_{\mathbf{p}, \mathbf{p}'} \quad \text{et} \quad \sum_{\mathbf{p}} |\mathbf{p}\rangle \langle \mathbf{p}| = 1 \quad (\text{ND}) \quad (1.51)$$

La fonction d'onde associée est

$$\langle \mathbf{r} | \mathbf{p} \rangle = \frac{1}{\sqrt{\mathcal{V}}} e^{i\mathbf{p}\cdot\mathbf{r}} \quad (1.52)$$

1.D.2 Normalisation continue des états

L'avantage de la normalisation discrète des états est sa clarté du point de vue du décompte des états. Son désavantage est la référence au volume \mathcal{V} de l'espace physique, qui est artificiel (la limite $\mathcal{V} \rightarrow \infty$ doit être prise à la fin des calculs, et le résultat final ne doit pas dépendre de \mathcal{V}). Un autre désavantage de cette normalisation est qu'elle n'est pas invariante de Lorentz : en effet, le volume est contracté d'un facteur $\gamma = 1/\sqrt{1-\beta^2}$ lorsqu'on passe d'un référentiel où le volume est au repos, à un référentiel où il est en mouvement à une vitesse β . Comme l'invariance de Lorentz est une symétrie fondamentale de la nature, particulièrement visible dans le domaine de particules élémentaires, il est souhaitable d'adopter une autre normalisation des états d'impulsion, sans référence à un volume fini de l'espace.

À cette fin, notons premièrement que la somme sur les vecteurs d'onde peut être remplacée par une intégrale, dans la limite où le volume \mathcal{V} est grand :

$$\sum_{\mathbf{p}} \rightarrow \mathcal{V} \int \frac{d^3 p}{(2\pi)^3} \quad (1.53)$$

Ceci provient directement de (1.49). En conséquence, le delta de Kronecker $\delta_{\mathbf{p}, \mathbf{p}'}$ peut être remplacé par une fonction delta de Dirac :

$$\delta_{\mathbf{p}, \mathbf{p}'} \rightarrow \frac{(2\pi)^3}{\mathcal{V}} \delta^3(\mathbf{p} - \mathbf{p}') \quad (1.54)$$

Il suffit, pour s'en convaincre, de constater l'équivalence de

$$\sum_{\mathbf{p}} \delta_{\mathbf{p}, \mathbf{p}'} \quad \text{avec} \quad \mathcal{V} \int \frac{d^3 p}{(2\pi)^3} \frac{(2\pi)^3}{\mathcal{V}} \delta^3(\mathbf{p} - \mathbf{p}') \quad (1.55)$$

Une première modification à la normalisation ND serait de supprimer le facteur de volume, c'est-à-dire de considérer des états d'impulsion dans le spectre continu tels que

$$\langle \mathbf{p}' | \mathbf{p} \rangle = (2\pi)^3 \delta^3(\mathbf{p} - \mathbf{p}') \quad \text{et} \quad \int \frac{d^3 p}{(2\pi)^3} |\mathbf{p}\rangle \langle \mathbf{p}| = 1 \quad (\text{NC}) \quad (1.56)$$

La fonction d'onde associée est alors

$$\langle \mathbf{r} | \mathbf{p} \rangle = e^{i\mathbf{p}\cdot\mathbf{r}} \quad (1.57)$$

Cependant, cette normalisation continue (appelons-la NC) n'est toujours pas invariante de Lorentz. En effet, la mesure d'intégration $d^3 p$ n'est pas invariante : il s'agit d'un élément de volume dans l'espace des impulsions. Lors d'un changement de référentiel, cet élément de volume est contracté d'un facteur γ .

1.D.3 Normalisation relativiste des états

Un élément de volume dans l'espace des quadri-impulsions

$$d^4p = d^3p \, dp^0 \quad (1.58)$$

où $p^0 = E$ est l'énergie de la particule, est invariant. En effet, la transformation de Lorentz implique une contraction des longueurs d'un facteur γ et une dilatation du temps d'un même facteur, de sorte que l'élément de volume d'espace-temps $d^3r \, dt$ est invariant. Il en est de même pour tout quadrivecteur, en particulier pour l'élément de volume de la quadri-impulsion. Nous ne pouvons cependant pas remplacer d^3p par d^4p sans autre forme de procès, car l'énergie p^0 est déterminée par la quantité de mouvement \mathbf{p} , en vertu de la relation

$$(p^0)^2 - \mathbf{p}^2 = E^2 - \mathbf{p}^2 = m^2 \quad \text{ou} \quad E = \sqrt{\mathbf{p}^2 + m^2} \quad (1.59)$$

On impose donc cette relation par une fonction delta supplémentaire, c'est-à-dire qu'on adopte la mesure d'intégration suivante :

$$\int \frac{d^3p}{(2\pi)^3} \rightarrow \int \frac{d^4p}{(2\pi)^3} \delta((p^0)^2 - \mathbf{p}^2 - m^2) \theta(p^0) \quad (1.60)$$

où la fonction de Heaviside $\theta(p^0)$ s'assure qu'on ne conserve que la racine positive, soit $p^0 = +\sqrt{\mathbf{p}^2 + m^2}$. Notons que cette mesure d'intégration est manifestement invariante de Lorentz, car (i) la mesure d^4p est invariante, comme noté ci-haut ; (ii) l'argument de la fonction delta est un invariant de Lorentz et (iii) comme le quadrivecteur (p^0, \mathbf{p}) est toujours de genre temps, si $p^0 > 0$ dans un référentiel, il l'est dans tous les référentiels et donc $\theta(p^0)$ est un invariant. On montre dans ce qui suit que cette mesure peut aussi s'exprimer comme

$$\int \frac{d^4p}{(2\pi)^3} \delta((p^0)^2 - \mathbf{p}^2 - m^2) \theta(p^0) = \int \frac{d^3p}{(2\pi)^3 2E} \quad (1.61)$$

En effet, la fonction delta peut s'écrire comme

$$\delta((p^0)^2 - \mathbf{p}^2 - m^2) = \delta((p^0 - E)(p^0 + E)) \quad \text{où} \quad E = \sqrt{\mathbf{p}^2 + m^2} \quad (1.62)$$

En intégrant sur p^0 , seule la racine $p^0 = E$ est considérée en raison du facteur $\theta(p^0)$. Au voisinage de cette racine, on peut remplacer $p^0 + E$ par $2E$ et la fonction delta se réduit à

$$\delta((p^0 - E)(p^0 + E)) \rightarrow \delta(2E(p^0 - E)) = \frac{1}{2E} \delta(p^0 - E) \quad (1.63)$$

On intègre ensuite sur p^0 pour retrouver le résultat annoncé, en remplaçant partout p^0 par E .

Une normalisation relativiste des états d'impulsion est donc fixée par les conditions

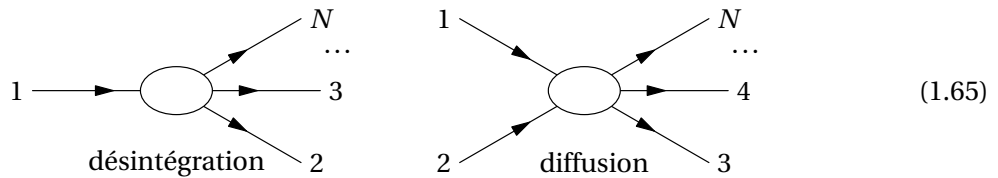
$$\langle \mathbf{p}' | \mathbf{p} \rangle = 2E (2\pi)^3 \delta^3(\mathbf{p} - \mathbf{p}') \quad \text{et} \quad \int \frac{d^3p}{(2\pi)^3 2E} |\mathbf{p}\rangle \langle \mathbf{p}| = 1 \quad (\text{NR}) \quad (1.64)$$

où $E = \sqrt{\mathbf{p}^2 + m^2}$.

E Processus de désintégration

Notre connaissance empirique des particules élémentaires provient d'expériences impliquant des collisions et/ou des désintégrations d'objets instables. Ces processus respectent tous les lois de conservations élémentaires de l'énergie et de la quantité de mouvement et se produisent généralement à des énergies assez élevées pour qu'une description relativiste des processus en cause soit nécessaire.

Les processus de désintégration et de collision (ou diffusion) sont illustrés schématiquement ci-dessous :



Les particules (ou objets) en jeu sont numérotées de 1 à N , et le temps s'écoule de gauche à droite. Une désintégration implique au moins trois objets : la particule instable (1) et au moins deux produits de désintégration (numérotés de 2 à N). On parle d'une désintégration à deux corps, à trois corps, etc., selon le nombre de produits. Une collision implique au moins quatre objets : deux réactants (1 et 2) et au moins deux produits (numérotés de 3 à N).

Ces processus sont un passage d'un état quantique initial $|i\rangle$ vers un état quantique final $|f\rangle$ différent. On suppose que les particules en cause, à l'état libre, sont décrites par un certain hamiltonien H_0 décrivant des particules libres. Si H_0 était le hamiltonien total du système, alors les particules seraient stables et aucune interaction (ou collision) ne serait possible. Les processus de désintégration et de collision sont donc attribuables à un hamiltonien supplémentaire, noté V et qualifié d'*interaction*, qui pousse les particules à se transformer : on décompose donc le hamiltonien total ainsi : $H = H_0 + V$.

Par exemple, si une particule se désintègre, c'est que son état de particule n'est pas un état propre du hamiltonien total H , mais de H_0 seulement. Ainsi, un état excité de l'atome d'hydrogène est un état propre du hamiltonien de l'atome d'hydrogène lui-même, mais n'est pas un état propre du hamiltonien complet qui comprend l'interaction entre l'atome et le champ électromagnétique. C'est donc pour cela que les états excités sont instables, et qu'ils se « désintègrent » en un atome d'hydrogène à l'état fondamental et un photon.

1.E.1 Règle d'or de Fermi

Si, dans un certain sens, le terme d'interaction V est « petit » devant H_0 , alors il ne se manifeste que par des processus, comme la désintégration et les collisions, qui se produisent suffisamment rarement pour que les particules isolées soient concevables. Autrement, les états de particules libres seraient si éloignés des véritables états propres du hamiltonien, que la notion même de particule

n'aurait pas beaucoup de sens. Il est alors raisonnable de traiter V comme une perturbation, d'autant plus que la résolution du hamiltonien complet est impossible. On applique alors la théorie des perturbations dépendantes du temps, via la fameuse règle d'or de Fermi :

$$\omega_{i \rightarrow f} = 2\pi |M_{fi}|^2 \delta(E_f - E_i) \quad (1.66)$$

où $\omega_{i \rightarrow f}$ est la probabilité par unité de temps (ou taux de transition) pour que l'état initial $|i\rangle$ se transmute en état final $|f\rangle$. L'énergie étant conservée, la fonction delta $\delta(E_i - E_f)$ nous assure que les énergies des deux états sont les mêmes. Enfin, M_{fi} est l'amplitude de transition entre les états initial et final, donnée au premier ordre de la théorie des perturbations par l'élément de matrice de la perturbation V entre ces deux états :

$$M_{fi} = \langle f | V | i \rangle \quad (1.67)$$

Il s'agit bien sûr d'une approximation, au premier ordre en V . Aux ordres suivants, l'amplitude comporte en plus une sommation sur des états intermédiaires :

$$M_{fi} = \langle f | V | i \rangle + \sum_n \frac{\langle f | V | n \rangle \langle n | V | i \rangle}{E_i - E_n + i0^+} + \sum_{n,m} \frac{\langle f | V | n \rangle \langle n | V | m \rangle \langle m | V | i \rangle}{(E_i - E_n + i0^+)(E_i - E_m + i0^+)} + \dots \quad (1.68)$$

Le deuxième terme est une somme sur les états propres $|n\rangle$ du hamiltonien non perturbé H_0 , dont les énergies sont E_n . Le troisième terme comporte deux sommations sur des états intermédiaires, et ainsi de suite. La partie imaginaire infinitésimale au dénominateur sert parfois à résoudre des ambiguïtés analytiques lors de la sommation sur les états intermédiaires. Le deuxième terme doit être considéré si le premier s'annule, ou afin d'améliorer la précision du calcul (nous allons nous limiter au premier terme dans ce qui suit, mais le deuxième ordre interviendra dans un chapitre ultérieur, lors de l'introduction des diagrammes de Feynman).

1.E.2 Forme générale du taux de désintégration

Lors d'une désintégration, il existe un très grand nombre (en pratique, un continuum) d'états finaux possibles. Ces états ont la même énergie que l'état initial, mais sont caractérisés par des impulsions différentes des particules émises. La quantité pertinente est la probabilité totale par unité de temps pour qu'une transition se produise vers l'un des états finaux appartenant à un ensemble expérimentalement identifiable.

Considérons un processus dans lequel une particule de masse m_1 se désintègre en $N-1$ particules, de masses m_i ($i = 2, \dots, N$). Les énergies des états initial et final sont $E_i = E_1$ et $E_f = E_2 + \dots + E_N$, respectivement. Le taux de transition de l'état initial vers un ensemble d'états finaux compris dans un élément d'espace des phases $d^3p_2 d^3p_3 \dots d^3p_N$ est (nous adoptons la normalisation NC des états d'impulsions, éq. (1.56))

$$d\omega = 2\pi |M_{fi}|^2 \frac{d^3p_2}{(2\pi)^3} \frac{d^3p_3}{(2\pi)^3} \dots \frac{d^3p_N}{(2\pi)^3} \delta(E_1 - E_2 - E_3 - \dots - E_N) \quad (1.69)$$

En général, le hamiltonien d'interaction V conserve la quantité de mouvement totale, de sorte que l'amplitude de transition n'est non nulle que si les impulsions respectent la condition $\mathbf{p}_1 = \mathbf{p}_2 + \mathbf{p}_3 + \dots + \mathbf{p}_N$. On peut donc écrire l'amplitude comme

$$M_{fi} = \mathcal{M}_{fi} \delta_{\mathbf{p}_1 - \mathbf{p}_2 - \mathbf{p}_3 - \dots - \mathbf{p}_N} \quad (1.70)$$

dans la normalisation discrète ND, éq. (1.51). En fait, cette relation définit \mathcal{M}_{fi} , qu'on appelle aussi *amplitude de transition*. Au carré, cette relation est simplement

$$|M_{fi}|^2 = |\mathcal{M}_{fi}|^2 \delta_{\mathbf{p}_1 - \mathbf{p}_2 - \mathbf{p}_3 - \dots - \mathbf{p}_N} \quad (1.71)$$

car le carré du delta de Kronecker est encore un delta de Kronecker. Notons ici que la normalisation discrète des états (ND) est la plus sûre quand vient le temps de compter les états ou de calculer le carré d'un delta de Kronecker. Par contre, quand vient le temps d'effectuer un calcul pratique, on lui préfère la normalisation continue NC, éq. (1.56). Avec cette nouvelle définition de l'amplitude, le taux de désintégration devient

$$d\omega = |\mathcal{M}_{fi}|^2 \frac{d^3 p_2}{(2\pi)^3} \frac{d^3 p_3}{(2\pi)^3} \dots \frac{d^3 p_N}{(2\pi)^3} (2\pi)^4 \delta^4(p_1 - p_2 - p_3 - \dots - p_N) \quad (1.72)$$

où p_i dénote le quadrivecteur (E_i, \mathbf{p}_i) .

Pour convertir cette expression en fonction de la normalisation relativiste (1.64), on doit simplement effectuer le remplacement

$$|\mathbf{p}_1\rangle_{\text{NC}} = \frac{1}{\sqrt{2E_1}} |\mathbf{p}_1\rangle_{\text{NR}} \quad (1.73)$$

dans l'amplitude de transition, ce qui donne

$$\mathcal{M}_{fi}|_{\text{NC}} = \frac{1}{\sqrt{2^N E_1 E_2 \dots E_N}} \mathcal{M}_{fi}|_{\text{NR}} \quad (1.74)$$

On obtient donc le résultat suivant (nous laissons tomber l'indice NR de l'amplitude) :

$$d\omega = |\mathcal{M}_{fi}|^2 \frac{1}{2E_1} \frac{d^3 p_2}{(2\pi)^3 2E_2} \frac{d^3 p_3}{(2\pi)^3 2E_3} \dots \frac{d^3 p_N}{(2\pi)^3 2E_N} (2\pi)^4 \delta^4(p_1 - p_2 - p_3 - \dots - p_N) \quad (1.75)$$

L'amplitude est dans cette normalisation une quantité invariante de Lorentz, qui a les dimensions (unités) de l'énergie dans le cas d'une désintégration à deux corps ($N = 3$).

Exemple 1.1 Désintégration à deux corps

Appliquons le résultat général (1.75) au cas d'une désintégration à deux corps. La formule est, dans ce cas,

$$d\omega = |\mathcal{M}_{fi}|^2 \frac{1}{2E_1} \frac{d^3 p_2}{(2\pi)^3 2E_2} \frac{d^3 p_3}{(2\pi)^3 2E_3} (2\pi)^4 \delta^4(p_1 - p_2 - p_3) \quad (1.76)$$

où $E_2 = \sqrt{\mathbf{p}_2^2 + m_2^2}$ et $E_3 = \sqrt{\mathbf{p}_3^2 + m_3^2}$. Plaçons-nous dans le référentiel de la particule instable, de sorte que $E_1 = m_1$ et $\mathbf{p}_1 = 0$. Le taux de désintégration total est obtenu en intégrant sur les états finaux possibles :

$$\omega = \int \frac{d^3 p_2}{(2\pi)^3 2E_2} \frac{d^3 p_3}{(2\pi)^3 2E_3} \frac{|\mathcal{M}_{fi}|^2}{2E_1} (2\pi)^4 \delta^4(p_1 - p_2 - p_3) \quad (1.77)$$

Utilisons la fonction delta pour intégrer sur \mathbf{p}_3 :

$$\omega = \int \frac{d^3 p_2}{(2\pi)^3} \frac{|\mathcal{M}_{fi}|^2}{8m_1} \frac{(2\pi) \delta(m_1 - \sqrt{p^2 + m_2^2} - \sqrt{p^2 + m_3^2})}{\sqrt{p^2 + m_2^2} \sqrt{p^2 + m_3^2}} \quad (1.78)$$

où on a utilisé le fait que $\mathbf{p}_3^2 = \mathbf{p}_2^2$ et où on pose simplement $p = |\mathbf{p}_2|$. Il est clair, par symétrie, que l'amplitude ne peut pas dépendre de la direction de \mathbf{p}_2 . On peut donc intégrer sur les angles et on trouve

$$\omega = \int p^2 dp \frac{|\mathcal{M}_{fi}|^2}{8\pi m_1} \frac{\delta(m_1 - \sqrt{p^2 + m_2^2} - \sqrt{p^2 + m_3^2})}{\sqrt{p^2 + m_2^2} \sqrt{p^2 + m_3^2}} \quad (1.79)$$

La dernière intégrale se fait facilement si on procède au changement de variable suivant : on pose

$$E = \sqrt{p^2 + m_2^2} + \sqrt{p^2 + m_3^2} \quad (1.80)$$

alors

$$\frac{dE}{dp} = \frac{p}{\sqrt{p^2 + m_2^2}} + \frac{p}{\sqrt{p^2 + m_3^2}} = \frac{Ep}{\sqrt{p^2 + m_2^2} \sqrt{p^2 + m_3^2}} \quad (1.81)$$

On trouve ensuite

$$\omega = \int_{m_2+m_3}^{\infty} dE p \frac{|\mathcal{M}_{fi}|^2}{8\pi m_1 E} \delta(m_1 - E) = |\mathbf{p}_2| \frac{|\mathcal{M}_{fi}|^2}{8\pi m_1^2} \quad (1.82)$$

en autant que les impulsions des deux produits permettent à leur énergie totale E d'être égale à m , c'est-à-dire que la désintégration soit énergétiquement possible. Dans cette formule $|\mathbf{p}_2|$ est la quantité de mouvement de l'un ou l'autre des produits de la désintégration, dans le référentiel de la particule instable. Ce résultat est très simple et pourtant très général, car nous n'avons pas besoin de connaître la forme détaillée de l'amplitude \mathcal{M}_{fi} pour le démontrer.

1.E.3 Loi exponentielle de désintégration

Le processus de désintégration est caractérisé principalement par le taux de désintégration décrit noté ω ci-dessus ou, plus souvent, noté Γ . Une probabilité constante de transition par unité de temps mène directement à une loi exponentielle de désintégration : Si $N(t)$ représente la population de particules se désintégrant à un instant t , alors le nombre de désintégrations ayant lieu entre les temps t et $t + dt$ est $N(t)\Gamma dt$. Donc on trouve l'équation différentielle suivante :

$$N(t + dt) = N(t) - \Gamma N dt \implies \frac{dN}{dt} = -\Gamma N \quad (1.83)$$

dont la solution est la loi exponentielle

$$N(t) = N(0)e^{-\Gamma t} \quad (1.84)$$

où $N(0)$ est le nombre de particules au temps $t = 0$.

Quelques définitions :

vie moyenne $\tau = 1/\Gamma$. Durée moyenne de la vie d'une particule.

Demi-vie ou **période** $t_{1/2} = \tau \ln 2$, temps au bout duquel la moitié de la population initiale s'est désintégrée.

Remarques :

- ◆ La loi exponentielle est statistique. Il s'agit de probabilités. Lorsque la population de particules n'est pas très grande, les fluctuations peuvent être importantes. Ces fluctuations sont gouvernées par la loi de Poisson (voir l'annexe 1.E.5).
- ◆ L'*activité* I (ou intensité) d'une source radioactive est définie simplement comme le nombre de désintégrations par unité de temps. Dans le cas d'un échantillon ne contenant qu'une espèce de particule instable, $I = \Gamma N = -dN/dt$. L'activité se mesure fondamentalement en secondes inverses : un *Becquerel* (Bq) est l'activité d'une source qui produit une désintégration par seconde. Un *Curie* (Ci) est l'activité d'un gramme de radium pur, soit $3,7 \times 10^{10}$ Bq.
- ◆ La loi exponentielle suppose qu'aucun processus de régénération n'existe. En général, un noyau instable est lui-même le fruit de la désintégration d'un autre noyau. Les équations de population d'espèces sont dans ce cas plus complexes, mais impliquent toujours des fonctions exponentielles. Par exemple, si un noyau 1 se désintègre en un noyau 2 avec un taux Γ_1 , et que le noyau 2 se désintègre en autre chose avec un taux Γ_2 , les populations N_1 et N_2 des deux espèces seraient régies par les équations différentielles suivantes :

$$\begin{aligned} \frac{dN_1}{dt} &= -\Gamma_1 N_1 \\ \frac{dN_2}{dt} &= -\Gamma_2 N_2 + \frac{dN_1}{dt} = -\Gamma_2 N_2 + \Gamma_1 N_1 \end{aligned} \quad (1.85)$$

et des modèles plus complexes sont possibles.

- ◆ La loi exponentielle est étroitement liée à l'existence du continuum d'états finaux $|f\rangle$. S'il n'y avait qu'un seul état final de même énergie que l'état initial, il n'y aurait pas décroissance exponentielle de l'état initial, mais plutôt oscillation entre les deux états (les oscillations de Rabi). La notion subtile de *décohérence* intervient également.
- ◆ Une particule ou noyau instable peut avoir plusieurs modes de désintégration, c'est-à-dire plusieurs types de produits. Dans ce cas, chaque mode est caractérisé par son propre taux de désintégration. Par exemple, si deux modes sont possibles, deux taux (Γ_1 et Γ_2) existent et le taux total de désintégration en est la somme : $\Gamma = \Gamma_1 + \Gamma_2$. Dans un intervalle de temps donné, le rapport des nombres N_1 et N_2 de désintégrations des deux types est $N_1/N_2 = \Gamma_1/\Gamma_2$.

1.E.4 Relation d'incertitude temps-énergie

Une particule instable, par définition, n'est pas dans un état propre de l'énergie. Une mesure de l'énergie de cette particule sera donc entachée d'une incertitude. Nous allons montrer dans cette section que cette incertitude ΔE est reliée au temps de vie τ par la relation

$$\tau \Delta E \sim 1 \quad (1.86)$$

Commençons par remarquer que si un état quantique $|\psi\rangle$ est instable, c'est que sa projection sur sa valeur initiale décroît exponentiellement. Autrement dit,

$$\langle \psi(0) | \psi(t) \rangle = e^{-iE_0 t} e^{-\Gamma t/2} \quad (1.87)$$

où E_0 est l'énergie approximative de l'état et Γ le taux de désintégration. En effet, la probabilité de trouver le système au temps t dans le même état où il était au temps $t = 0$ est

$$|\langle \psi(0) | \psi(t) \rangle|^2 = e^{-\Gamma t} \quad (1.88)$$

ce qui correspond bien à la notion de vie moyenne $\tau = \Gamma^{-1}$.

Introduisons ici le concept de *fonction spectrale* $A(E)$ associée à l'état $|\psi\rangle$. Si on désigne par $|n\rangle$ les véritables états stationnaires du système physique étudié (donc les états propres du hamiltonien complet), la fonction spectrale $A(E)$ est la probabilité qu'une mesure de l'énergie dans l'état $|\psi\rangle$ donne E :

$$A(E) = \langle \psi | \delta(E - H) | \psi \rangle = \sum_n \delta(E - E_n) |\langle \psi | n \rangle|^2 \quad (1.89)$$

La fonction spectrale est reliée à la *résolvante* $G(z)$, définie comme suit :

$$G(z) = \langle \psi | \frac{1}{z - H} | \psi \rangle \quad (1.90)$$

où z est une variable complexe. En effet, on montre sans peine que

$$A(E) = -\frac{1}{\pi} \lim_{\eta \rightarrow 0^+} \text{Im} G(E + i\eta) \quad (1.91)$$

Il suffit pour cela d'exprimer la résolvante en fonction de la base des états $|n\rangle$:

$$G(z) = \sum_n |\langle \psi | n \rangle|^2 \frac{1}{z - E_n} \quad (1.92)$$

et de constater que

$$-\frac{1}{\pi} \text{Im} \frac{1}{E - E_n + i\eta} = \frac{1}{\pi} \frac{\eta}{(E - E_n)^2 + \eta^2} \quad (1.93)$$

Or, la fonction

$$\frac{1}{\pi} \frac{\eta}{x^2 + \eta^2} \quad (1.94)$$

tend vers la fonction $\delta(x)$ quand $\eta \rightarrow 0^+$. Donc, dans cette limite, on retrouve bien la fonction spectrale.

D'autre part, si on calcule la *fonction d'autocorrélation*

$$S(E) = \int_0^\infty dt e^{iEt - \eta t} \langle \psi(0) | \psi(t) \rangle \quad (\eta \rightarrow 0^+) \quad (1.95)$$

on trouve

$$S(E) = \sum_n \frac{i}{E - E_n + i\eta} |\langle \psi | n \rangle|^2 = iG(E + i\eta) \quad (1.96)$$

Donc on peut retrouver la fonction spectrale $A(E)$ en calculant $S(E)$. Pour un état en décroissance exponentielle, on a justement

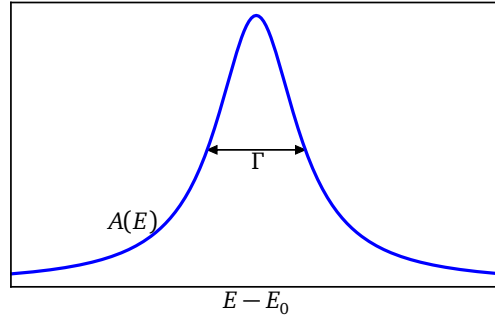
$$S(E) = \int_0^\infty dt e^{-iE_0 t - \Gamma t/2} e^{iEt - \eta t} = \frac{i}{E - E_0 + i\Gamma/2} \quad (1.97)$$

Notons que nous avons pu prendre la limite $\eta \rightarrow 0$ dès à présent, car Γ joue le même rôle que η dans l'intégration, et donc le facteur de convergence $e^{-\eta t}$ n'est plus nécessaire. Ceci correspond à une résolvante

$$G(z) = \frac{1}{z - E_0 + i\Gamma/2} \quad (1.98)$$

et donc à une fonction spectrale

$$A(E) = \frac{1}{2\pi} \frac{\Gamma}{(E - E_0)^2 + (\Gamma/2)^2} \quad (1.99)$$



On constate que Γ est la largeur à mi-hauteur de la lorentzienne : $A(E_0 + \Gamma/2) = \frac{1}{2}A(E_0)$.

Résumons : si on effectue une mesure de l'énergie du système, on ne trouvera pas toujours l'énergie approximative E_0 , mais plutôt une distribution de probabilité de l'énergie donnée par une courbe lorentzienne dont la largeur à mi-hauteur est le taux de désintégration Γ de l'objet considéré. Il arrive souvent que la vie moyenne d'une particule instable soit mesurée de cette manière, en accumulant des statistiques sur des mesures de son énergie. En pratique, il y a aussi une erreur expérimentale (liée à la précision des instruments) sur l'énergie : on parle alors de *largeur extrinsèque*, s'appliquant généralement à une forme gaussienne. Par contre, si la résolution de l'instrument est suffisante, on peut avoir accès à la *largeur intrinsèque* Γ .

1.E.5 Annexe : distribution de Poisson

Nous allons montrer dans cette annexe que si une désintégration se produit en moyenne λ fois dans un intervalle de temps donné, alors la probabilité qu'elle se produise k fois dans cet intervalle est donnée par la distribution de Poisson

$$P(k) = \frac{\lambda^k e^{-\lambda}}{k!} \quad (1.100)$$

En outre, nous allons montrer que la variance du nombre k de désintégrations dans ce même intervalle de temps est $\sigma_k^2 = \lambda$.

Commençons par diviser l'intervalle de temps en n sous-intervalles égaux, où n est suffisamment grand pour qu'on puisse négliger la possibilité que deux désintégrations se produisent à l'intérieur d'un même sous-intervalle. La probabilité qu'une désintégration se produise dans un sous-intervalle est alors λ/n , et la probabilité que k désintégrations se produisent dans l'intervalle to-

tal est donné par la distribution binomiale, puisque les différentes désintégrations sont des événements indépendants :

$$B(n, k) = \frac{n!}{k!(n-k)!} \left(\frac{\lambda}{n}\right)^k \left(1 - \frac{\lambda}{n}\right)^{n-k} \quad (1.101)$$

Cette probabilité est le nombre de combinaisons possibles de k sous-intervalles avec désintégration parmi n sous-intervalles possibles, fois la probabilité que ces k sous-intervalles contiennent une désintégration et que les $n - k$ autres n'en contiennent pas.

Il faut ensuite prendre la limite $n \rightarrow \infty$, pour respecter la condition essentielle qu'un même sous-intervalle ne puisse contenir plus qu'une désintégration. Or, dans cette limite,

$$\frac{n!}{k!(n-k)!} = \frac{n(n-1)(n-2)\cdots(n-k+1)}{k!} \rightarrow \frac{n^k}{k!} \quad (1.102)$$

alors que

$$\left(1 - \frac{\lambda}{n}\right)^{n-k} = \left(1 - \frac{\lambda}{n}\right)^n \left(1 - \frac{\lambda}{n}\right)^{-k} \rightarrow e^{-\lambda} \cdot 1 \quad (1.103)$$

Donc au total

$$\lim_{n \rightarrow \infty} B(n, k) = \lim_{n \rightarrow \infty} \frac{n^k}{k!} \left(\frac{\lambda}{n}\right)^k e^{-\lambda} = \frac{\lambda^k e^{-\lambda}}{k!} \quad (1.104)$$

et nous obtenons bel et bien la distribution de Poisson (1.100).

Calculons maintenant les premiers moments de cette distribution. Pour ce faire, considérons la fonction génératrice des moments :

$$M(t) := \sum_{k=0}^{\infty} e^{tk} P(k) = \sum_{k=0}^{\infty} e^{tk} \frac{\lambda^k e^{-\lambda}}{k!} = \sum_{k=0}^{\infty} \frac{(e^t \lambda)^k e^{-\lambda}}{k!} = e^{-\lambda} e^{e^t \lambda} = e^{\lambda(e^t - 1)} \quad (1.105)$$

D'après la définition de cette fonction génératrice, la normalisation de la distribution est $M(0)$, sa moyenne est $M'(0)$ et la moyenne du carré de k est $M''(0)$, etc. Or, on vérifie que

$$M(0) = 1 \quad M'(t) = \lambda e^t e^{\lambda(e^t - 1)} \quad M''(t) = (\lambda e^t + [\lambda e^t]^2) e^{\lambda(e^t - 1)} \quad (1.106)$$

et donc $\langle k \rangle = M'(0) = \lambda$, alors que la variance est

$$\sigma_k^2 = M''(0) - [M'(0)]^2 = (\lambda + \lambda^2) - \lambda^2 = \lambda \quad (1.107)$$

l'écart-type de la distribution de Poisson est donc $\sqrt{\lambda}$.

F Collisions et section efficace

1.F.1 Notion de section efficace

La notion de *section efficace* est sans doute la plus importante de toute la physique des particules expérimentale. Considérons un faisceau de particules, ayant toutes la même vitesse et la même direction, se dirigeant vers une cible. Ce faisceau est caractérisé par un flux Φ , soit le nombre de particules par unité de temps et de surface traversant une surface perpendiculaire au faisceau. Chaque particule peut interagir avec un objet dans la cible, le *diffuseur*, et être déviée dans une direction particulière (θ, φ) (adoptons un système de coordonnées sphériques, avec l'axe du faisceau comme axe des z). On supposera, pour simplifier, que le diffuseur est infiniment massif en comparaison des particules du faisceau et qu'il n'est pas affecté par les collisions. En principe, la direction de diffusion (θ, φ) est uniquement déterminée par la position (x, y) de la particule incidente par rapport à l'axe z (le paramètre d'impact) et par la forme précise de son interaction avec l'objet diffuseur (comme le potentiel d'interaction).

La *section différentielle de diffusion* est alors définie comme le nombre de particules diffusées dans la direction (θ, φ) , par unité de temps et d'angle solide, divisé par le flux incident Φ :

$$\frac{d\sigma}{d\Omega} = \frac{\# \text{ de particules diffusées vers } (\theta, \varphi) \text{ par sec.}}{\text{flux incident} \times \text{angle solide } d\Omega} \quad (1.108)$$

La section différentielle est une fonction des angles (θ, φ) et possède les unités d'une surface. La *section efficace* σ est alors le nombre de particules déviées par unité de temps, divisé par le flux incident, soit l'intégrale sur les angles de la section différentielle :

$$\sigma = \int d\Omega \frac{d\sigma}{d\Omega} \quad (1.109)$$

Autrement dit, le nombre de particules déviées par unité de temps est $\sigma\Phi$.

Dans le cas d'une interaction de contact entre les particules ponctuelles et le diffuseur, ces dernières ne sont diffusées que si elles entrent en contact direct avec l'objet. Le nombre de particules dans cette situation (par unité de temps) est précisément Φ multiplié par l'aire transversale A de l'objet. Donc, dans ce cas, on trouve $\sigma = A$, d'où le nom de *section efficace*. En somme, dans un problème plus général (sans interaction de contact), la section efficace nous indique la capacité d'un diffuseur à dévier les particules incidentes, en donnant la superficie équivalente d'un objet qui diffuserait uniquement par contact.

Longueur d'atténuation Considérons un milieu comportant une densité ρ de cibles (nombre par unité de volume), ainsi qu'un faisceau de projectiles incident sur ce milieu. Après avoir traversé une épaisseur dx du milieu, la fraction de projectiles diffusés sera $\rho\sigma dx$. On peut le voir comme suit : soit une aire A transversale au flux incident. Dans le volume délimité par cette aire et par l'épaisseur dx il y a $\rho A dx$ cibles. La probabilité qu'un projectile soit diffusé par

une cible en particulier en passant dans cette aire A est σ/A . Donc la probabilité qu'il soit diffusé par l'une des cibles présentes est

$$\rho A dx \times \frac{\sigma}{A} = \rho \sigma dx \quad (1.110)$$

Cela suppose bien sûr que les différentes cibles diffusent de manière incohérente; autrement dit, que les ondes de matière diffusées sur des cibles différentes n'interfèrent pas entre elles. Dans ce cas, le flux $\Phi(x)$ du faisceau décroît en rapport avec la proportion de particules diffusées entre les positions x et $x + dx$:

$$\Phi(x + dx) = \Phi(x) - \Phi(x) \rho \sigma dx \implies \frac{d\Phi}{dx} = -\Phi \rho \sigma \quad (1.111)$$

Il s'ensuit que l'intensité du faisceau incident diminuera de façon exponentielle en fonction de x :

$$\Phi(x) = \Phi(0) e^{-\rho \sigma x} \quad (1.112)$$

On peut aussi décrire cette atténuation par une longueur caractéristique $\xi = (\rho \sigma)^{-1}$, appelée *longueur d'atténuation*.

1.F.2 Diffusion par un potentiel

Considérons le problème d'un projectile non relativiste de masse m qui entre en collision avec une cible qu'on suppose fixe. L'état initial du projectile est une onde plane d'impulsion \mathbf{p}_1 , et son état final, après la collision, est une onde plane d'impulsion \mathbf{p}_2 . Une quantité de mouvement $\mathbf{q} = \mathbf{p}_2 - \mathbf{p}_1$ a été transférée à la cible, mais nous allons négliger l'énergie cinétique associée, comme si la cible était infiniment massive. Il s'agit en somme d'un problème à un corps. Nous supposons que l'état de la cible ne change pas, de sorte que la collision est élastique et que l'effet de la cible peut être décrit par un potentiel diffuseur $V(\mathbf{r})$, qu'on traitera comme une perturbation ajoutée au hamiltonien décrivant des particules libres. Nous traiterons ce problème dans le régime non relativiste, en utilisant la théorie des perturbations, via la règle d'or de Fermi (1.66).

Somme sur les états finaux Comme il est impossible en pratique de résoudre un état quantique de vecteur d'onde donné, il nous faut sommer cette probabilité de transition sur un ensemble d'états finaux. Cette somme peut se faire en intégrant sur \mathbf{p}_2 , en utilisant la mesure d'intégration appropriée pour compter les états dans la normalisation utilisée. Utilisons ici la normalisation discrète des états (ND) décrite à la section 1.D. On suppose alors que le système étudié est contenu dans un grand domaine périodique de volume \mathcal{V} . Les vecteurs d'ondes possibles dans un tel domaine périodique sont discrets, et une somme sur ces vecteurs d'ondes peut être remplacée par une intégrale de la manière suivante :

$$\sum_{\mathbf{p}} \longrightarrow \mathcal{V} \int \frac{d^3 p}{(2\pi)^3} \quad (1.113)$$

Le taux de transition pour un ensemble quasi continu de vecteurs d'ondes \mathbf{p}_2 est donc

$$\omega = \frac{2\pi}{\mathcal{V}} \int \frac{d^3 p_2}{(2\pi)^3} |M_{fi}|^2 \delta(E_1 - E_2) \quad (1.114)$$

Cette expression est alors la probabilité par unité de temps que la particule transite vers un état final contenu dans le domaine d'intégration sur \mathbf{p}_2 . En pratique, la fonction delta va sélectionner, dans ce domaine d'intégration, les états qui ont la même énergie que la particule incidente. Transformons la mesure d'intégration afin de faire apparaître une différentielle d'angle solide, en passant en coordonnées polaires : $d^3 p_2 = d\Omega p_2^2 dp_2$ où $d\Omega = \sin \theta d\theta d\varphi = |d\cos \theta| d\varphi$. La probabilité de transition par unité de temps vers un ensemble d'états d'impulsions comprises dans le domaine d'intégration est donc

$$\omega = \frac{1}{(2\pi)^2} \mathcal{V} \int |M_{fi}|^2 p_2^2 dp_2 d\Omega \delta(E_1 - E_2) \quad (1.115)$$

Par unité d'angle solide, ce taux de transition est simplement

$$\frac{d\omega}{d\Omega} = \frac{1}{(2\pi)^2} \mathcal{V} \int |M_{fi}|^2 p_2^2 dp_2 \delta(E_1 - E_2) \quad (1.116)$$

où l'intégrale ne porte maintenant que sur le module p_2 . Comme $E_2 = p_2^2/2m$, on peut intégrer sur l'énergie au lieu de p_2 , sachant que $dE_2 = p_2 dp_2/m$. Ceci permet d'éliminer la fonction delta :

$$\frac{d\omega}{d\Omega} = \frac{\mathcal{V}}{(2\pi)^2} |M_{fi}|^2 p m \quad (1.117)$$

où maintenant $p = |\mathbf{p}_1| = |\mathbf{p}_2|$.

Cette quantité en soi n'est pas directement reliée à l'expérience, car en pratique on n'envoie pas un projectile à la fois, mais un faisceau de projectiles. On doit plutôt calculer la section différentielle de diffusion, obtenue en divisant ce taux de probabilité par le flux Φ de particules, égal à la densité de courant de projectiles, soit la vitesse v multipliée par la densité ρ . Si on utilise des états normalisés, il y a un seul projectile dans le domaine et la densité est $\rho = 1/\mathcal{V}$. La vitesse est égale à p/m , donc $\Phi = p/m \mathcal{V}$. La section différentielle de diffusion est alors

$$\frac{d\sigma}{d\Omega} = \frac{1}{\Phi} \frac{d\omega}{d\Omega} = \left(\frac{m}{2\pi} \right)^2 \mathcal{V}^2 |M_{fi}|^2 \quad (1.118)$$

Amplitude de diffusion Au premier ordre en théorie des perturbations, l'amplitude de diffusion est donnée par

$$\begin{aligned} M_{fi} &= \langle \mathbf{p}_2 | V | \mathbf{p}_1 \rangle = \int d^3 r \langle \mathbf{p}_2 | \mathbf{r} \rangle V(\mathbf{r}) \langle \mathbf{r} | \mathbf{p}_1 \rangle \\ &= \frac{1}{\mathcal{V}} \int d^3 r e^{-i(\mathbf{p}_2 - \mathbf{p}_1) \cdot \mathbf{r}} V(\mathbf{r}) \\ &= \frac{1}{\mathcal{V}} \tilde{V}(\mathbf{p}_2 - \mathbf{p}_1) \end{aligned} \quad (1.119)$$

où $\tilde{V}(\mathbf{q})$ désigne la transformée de Fourier du potentiel $V(\mathbf{r})$, évaluée au vecteur d'onde $\mathbf{q} = \mathbf{p}_2 - \mathbf{p}_1$, l'impulsion donnée au projectile par la cible. On trouve donc la section différentielle suivante :

$$\frac{d\sigma}{d\Omega} = \left(\frac{m}{2\pi} \right)^2 |\tilde{V}(\mathbf{q})|^2 \quad (\mathbf{q} = \mathbf{p}_2 - \mathbf{p}_1) \quad (1.120)$$

Notons que les facteurs de volume ont disparu ($\mathcal{V} M_{fi}$ est indépendant du volume).

Étudions le cas particulier de la collision entre une particule chargée (masse m , charge e_1) et une cible massive (masse $M \gg m$, charge e_2), sous l'effet de la seule force électrique entre les deux particules. Le potentiel associé est

$$V(\mathbf{r}) = \frac{e_1 e_2}{4\pi r} \quad (1.121)$$

Le problème, mathématiquement parlant, est de calculer la transformée de Fourier de ce potentiel.

On peut deviner la forme de $\tilde{V}(\mathbf{q})$ sans faire de calcul. Comme $V(\mathbf{r})$ ne dépend pas de la direction de \mathbf{r} , sa transformée de Fourier ne peut pas dépendre de la direction de \mathbf{q} et donc ne doit dépendre que de $|\mathbf{q}|$. D'autre part, les unités de $\tilde{V}(\mathbf{q})$ sont, d'après sa définition, celles d'une surface ($[L]^2$). Comme la transformation de Fourier est une opération linéaire, on obtient nécessairement la forme suivante :

$$\tilde{V}(\mathbf{q}) = \text{cte} \frac{e_1 e_2}{q^2} \quad (1.122)$$

où cte est une constante purement numérique. Nous allons montrer que cette constante est l'unité, de deux manières différentes.

Commençons par la manière courte. Nous savons que la fonction $\phi(r) = 1/4\pi r$ obéit à l'équation de Poisson :

$$\nabla^2 \phi = -\delta(\mathbf{r}) \quad (1.123)$$

En prenant la transformée de Fourier des deux membres de cette équation, on trouve

$$-\mathbf{q}^2 \tilde{\phi}(\mathbf{q}) = -1 \quad \text{et donc} \quad \tilde{\phi}(\mathbf{q}) = \frac{1}{\mathbf{q}^2}. \quad (1.124)$$

En effet, en calculant la transformée inverse,

$$\nabla^2 \phi = \nabla^2 \int \frac{d^3 q}{(2\pi)^3} e^{i\mathbf{q}\cdot\mathbf{r}} \tilde{\phi}(\mathbf{q}) = - \int \frac{d^3 q}{(2\pi)^3} e^{i\mathbf{q}\cdot\mathbf{r}} \mathbf{q}^2 \tilde{\phi}(\mathbf{q}) \quad (1.125)$$

et donc la transformée de Fourier de $\nabla^2 \phi$ est $-\mathbf{q}^2 \tilde{\phi}(\mathbf{q})$; d'autre part, la transformée de Fourier d'une fonction delta est l'unité. Il suffit alors de multiplier ce résultat par $e_1 e_2$ pour démontrer (1.122) avec cte = 1.

La manière longue consiste à calculer explicitement la transformée de Fourier. On procède comme suit, en coordonnées sphériques :

$$\begin{aligned} \tilde{V}(\mathbf{q}) &= \frac{e_1 e_2}{4\pi} \int r^2 dr \sin \theta d\theta d\varphi \frac{1}{r} e^{-iqr \cos \theta} \\ &= \frac{e_1 e_2}{2} \int_0^\infty dr r^2 \int_{-1}^1 dz \frac{e^{-iqzr}}{r} \quad (z = \cos \theta) \\ &= \frac{e_1 e_2}{q} \int_0^\infty dr \sin qr \end{aligned} \quad (1.126)$$

Malheureusement, cette dernière intégrale est mal définie. Ce problème est lié au fait que le potentiel de Coulomb a un rayon d'action infini et que sa section efficace est, strictement parlant, infinie.

Potentiel de Yukawa Pour remédier à cette situation, modifions le potentiel de Coulomb pour lui donner un rayon d'action fini (ce qu'on appelle le potentiel de Yukawa) :

$$\frac{e_1 e_2}{4\pi r} \rightarrow e_1 e_2 \frac{e^{-\mu r}}{4\pi r} \quad (1.127)$$

Ici μ^{-1} est une longueur caractéristique (on fera tendre μ vers 0 à la fin du calcul). Sa transformée de Fourier est alors

$$\begin{aligned} \tilde{V}(\mathbf{q}) &= \frac{e_1 e_2}{q} \int_0^\infty dr e^{-\mu r} \sin qr \\ &= \frac{e_1 e_2}{q} \int_0^\infty dr \frac{1}{2i} e^{-\mu r} (e^{iqr} - e^{-iqr}) \\ &= \frac{e_1 e_2}{q} \frac{1}{2i} \left(\frac{1}{\mu - iq} - \frac{1}{\mu + iq} \right) \\ &= \frac{e_1 e_2}{q^2 + \mu^2} \end{aligned} \quad (1.128)$$

On peut ensuite prendre la limite $\mu \rightarrow 0$ pour trouver

$$\tilde{V}(\mathbf{q}) = \frac{e_1 e_2}{q^2} \quad (1.129)$$

Expression de la section différentielle La section différentielle résultant de ce calcul est donc

$$\frac{d\sigma}{d\Omega} = \left(\frac{2m e_1 e_2}{4\pi q^2} \right)^2 \quad (1.130)$$

En posant $\mathbf{p}_1 = p\mathbf{z}$ et $\mathbf{p}_2 = p\mathbf{e}_r$ (coordonnées sphériques), on calcule que

$$q^2 = (\mathbf{p}_2 - \mathbf{p}_1)^2 = p^2(\mathbf{e}_r - \mathbf{z})^2 = p^2(2 - 2\cos\theta) = 4p^2 \sin^2 \theta / 2 \quad (1.131)$$

On peut donc écrire

$$\frac{d\sigma}{d\Omega} = \left(\frac{m e_1 e_2}{8\pi p^2} \right)^2 \text{cosec}^4 \frac{\theta}{2} \quad (1.132)$$

En fonction de l'énergie cinétique non relativiste $T = p^2/2m$, on retrouve bien la formule de Rutherford :

$$\frac{d\sigma}{d\Omega} = \left(\frac{e_1 e_2}{4\pi} \right)^2 \frac{1}{16T^2} \text{cosec}^4 \frac{\theta}{2} \quad (1.133)$$

On constate que la section efficace σ est infinie :

$$\sigma = \left(\frac{e_1 e_2}{16\pi T} \right)^2 \int d\varphi \sin\theta d\theta \text{cosec}^4 \frac{\theta}{2} \rightarrow \infty \quad (1.134)$$

en raison du comportement en θ^{-3} de l'intégrant au voisinage de $\theta = 0$. Ceci est encore dû au rayon d'action infini du potentiel de Coulomb. En pratique, le potentiel de Coulomb est écranté par d'autres charges et son rayon d'action efficace est fini, ce qui élimine ce problème. Par exemple, lors de la diffusion d'une particule α par un atome d'or, la charge électronique entourant le noyau rend l'atome effectivement neutre sur une distance d'environ 10^{-10}m , ce qui élimine le caractère non intégrable de la singularité à $\theta = 0$.

1.F.3 Forme relativiste générale de la section efficace

Considérons maintenant un processus de collision quelconque et obtenons une expression générale pour la section différentielle de diffusion. Considérons le processus illustré à droite de (1.65), où deux particules d'impulsions \mathbf{p}_1 et \mathbf{p}_2 se transforment en $N - 2$ autres particules d'impulsions $\mathbf{p}_3, \dots, \mathbf{p}_N$. Le taux de transition associé est, toujours dans la normalisation NC et par analogie avec la relation (1.72),

$$d\omega = |\mathcal{M}_{fi}|^2 \frac{d^3 p_3}{(2\pi)^3} \frac{d^3 p_4}{(2\pi)^3} \dots \frac{d^3 p_N}{(2\pi)^3} (2\pi)^4 \delta^4(p_1 + p_2 - p_3 - \dots p_N) \quad (1.135)$$

Plaçons-nous dans le référentiel du laboratoire, tel que $\mathbf{p}_2 = 0$. La section différentielle s'obtient, dans cette normalisation (NC) en divisant par la vitesse du projectile, soit par $|\mathbf{p}_1|/E_1$:

$$d\sigma = |\mathcal{M}_{fi}|^2 \frac{E_1}{|\mathbf{p}_1|} \frac{d^3 p_3}{(2\pi)^3} \frac{d^3 p_4}{(2\pi)^3} \dots \frac{d^3 p_N}{(2\pi)^3} (2\pi)^4 \delta^4(p_1 + p_2 - p_3 - \dots p_N) \quad (1.136)$$

Passons maintenant à la normalisation relativiste (1.64), ce qui se fait par le même changement de normalisation qu'à l'éq. (1.74) (dans ce cas $E_2 = m_2$) :

$$d\sigma = \frac{|\mathcal{M}_{fi}|^2}{4m_2|\mathbf{p}_1|} \frac{d^3 p_3}{(2\pi)^3 2E_3} \dots \frac{d^3 p_N}{(2\pi)^3 2E_N} (2\pi)^4 \delta^4(p_1 + p_2 - p_3 - \dots p_N) \quad (1.137)$$

Cette expression n'est valable que dans le référentiel du laboratoire. Pour la généraliser à un référentiel quelconque, on doit trouver une expression manifestement invariante de Lorentz qui se réduise à la forme ci-haut dans le repère du laboratoire. En effet, la section efficace est un invariant de Lorentz, car elle est une aire transversale à la direction de l'impulsion du projectile, et donc n'est pas affectée par une transformation de Lorentz effectuée dans cette direction. Vu différemment, la section efficace est une probabilité par unité de temps, divisée par un flux, qui est un nombre de particules par unité de temps et de surface. Le fait de procéder à une transformation de Lorentz dans la direction de l'impulsion apporte certainement un facteur de dilatation du temps, mais de manière égale au numérateur et au dénominateur, sans par ailleurs affecter les aires transversales à la quantité de mouvement. Dans l'expression (1.137), la mesure d'intégration est invariante de Lorentz, de même que l'amplitude (car nous avons adopté une normalisation des états qui est aussi invariante de Lorentz). Il reste l'expression $m_2 p_1$ du dénominateur qui ne l'est pas. Mais cette expression est un cas particulier dans le référentiel du laboratoire de l'expression invariante $\sqrt{(p_1 \cdot p_2)^2 - (m_1 m_2)^2}$. En effet, dans ce référentiel, $p_1 = (E_1, \mathbf{p}_1)$ et $p_2 = (m_2, 0)$, donc $p_1 \cdot p_2 = E_1 m_2$ et

$$(p_1 \cdot p_2)^2 - (m_1 m_2)^2 = E_1^2 m_2^2 - m_1^2 m_2^2 = (E_1^2 - m_1^2) m_2^2 = |\mathbf{p}_1|^2 m_2^2 \quad (1.138)$$

Donc $\sqrt{(p_1 \cdot p_2)^2 - (m_1 m_2)^2} \rightarrow |\mathbf{p}_1| m_2$. L'expression générale de la section différentielle est donc, dans la normalisation relativiste et dans un référentiel quelconque,

$$d\sigma = \frac{|\mathcal{M}_{fi}|^2}{4\sqrt{(p_1 \cdot p_2)^2 - (m_1 m_2)^2}} \frac{d^3 p_3}{(2\pi)^3 2E_3} \dots \frac{d^3 p_N}{(2\pi)^3 2E_N} (2\pi)^4 \delta^4(p_1 + p_2 - p_3 - \dots p_N) \quad (1.139)$$

1.F.4 Résonances et masse invariante

Le processus de collision de deux particules peut en principe mener à la création de nouvelles particules. L'apparition d'une particule de masse m est possible d'un point de vue énergétique si l'énergie totale des particules entrant en collision est au moins égale à m dans le référentiel du centre d'impulsion. En fait, quel que soit le référentiel utilisé, l'énergie disponible peut être obtenue en calculant la *masse invariante* associée aux impulsions et énergies des particules impliquées. Expliquons de quoi il s'agit. Si E_1 et E_2 désignent les énergies des deux particules, et \mathbf{p}_1 et \mathbf{p}_2 leurs impulsions, alors on appelle *masse invariante* M l'invariant associé au quadrivecteur $(E_1 + E_2, \mathbf{p}_1 + \mathbf{p}_2)$, soit

$$M^2 = (E_1 + E_2)^2 - (\mathbf{p}_1 + \mathbf{p}_2)^2 \quad (1.140)$$

Dans le référentiel du centre d'impulsion, $\mathbf{p}_1 + \mathbf{p}_2 = 0$ et la masse invariante est simplement la somme des énergies des particules, soit l'énergie disponible pour la création d'une particule de masse M (en supposant que les particules 1 et 2 sont annihilées). L'invariant M^2 est aussi noté s et appelé *première variable de Mandelstam*.

S'il est possible de créer une particule de masse m en annihilant deux particules de masses m_1 et m_2 , alors la section efficace σ , si on la considère comme une fonction de la masse invariante M , devrait comporter un maximum évident à $M = m$, car à cette énergie une nouvelle possibilité de transition quantique apparaît, qui n'existait pas à plus faible énergie. Cette possibilité s'ajoute à la possibilité de diffusion élastique, où les produits de la collision sont les particules initiales, quoique dans des états différents. D'autre part, la possibilité de création de cette nouvelle particule n'existe qu'à cette valeur de M et à aucune autre, car elle ne peut pas avoir d'autre énergie que m dans son propre référentiel. Donc, si cette particule était stable, un pic delta apparaîtrait dans la section efficace σ à une valeur de l'énergie incidente telle que la masse invariante M vaut m .

Or, la particule ainsi créée ne peut être stable. En effet, s'il est possible de la créer, il est également possible qu'elle se désintègre, par le processus inverse qui a mené à sa création, à savoir l'émission de deux particules. Cette particule possède donc une vie moyenne $\tau = \Gamma^{-1}$, Γ étant le taux de désintégration (parfois noté ω).

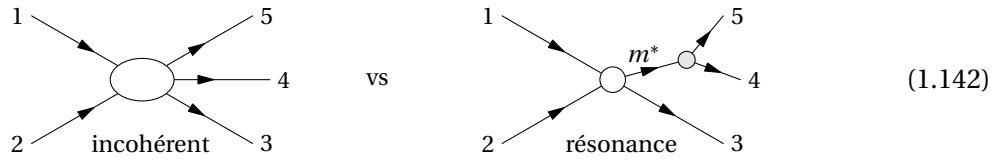
Par le principe d'incertitude temps-énergie, cela implique que son énergie dans son propre référentiel n'est pas bien définie, mais comporte une incertitude $\Delta E \sim \Gamma$. La densité d'états $\rho(E)$ associée n'est plus une fonction delta $\delta(E - m)$, mais plutôt une lorentzienne, comme justifié à la section 1.E.4 :

$$\rho(E) = \frac{1}{2\pi} \frac{\Gamma}{(E - m)^2 + (\Gamma/2)^2} \quad (1.141)$$

Il est donc possible, par une mesure de la largeur du pic lorentzien apparaissant dans la section efficace, de mesurer le temps de vie d'une particule créée lors d'un processus de collision. Ces considérations sont à la base de toutes les découvertes expérimentales de la physique des particules, mais s'appliquent aussi à des domaines en apparence aussi éloignés que la physique du solide, car les mêmes principes de base sont en jeu (exception faite de l'invariance relativiste).

La notion de masse invariante est aussi utilisée dans l'analyse des produits d'une réaction, afin de déterminer si une particule instable n'aurait pas été produite lors de la réaction, même si elle n'est

pas détectée dans les produits finaux en raison de sa désintégration subséquente. Par exemple, considérons une réaction où deux particules (1 et 2) donnent lieu à trois produits (3, 4 et 5) détectés :



Les techniques de détection nous donnent accès à la quantité de mouvement et à l'énergie des produits et on peut former la masse invariante $M = \sqrt{(E_4 + E_5)^2 - (\mathbf{p}_4 + \mathbf{p}_5)^2}$. Si la section efficace, portée en fonction de M , présente un pic significatif à une valeur précise (appelons-la m^*), c'est qu'on peut interpréter la réaction comme donnant naissance à deux particules, de masses m_3 et m^* . Cette réaction particulière ne peut se produire que si la masse invariante M est égale à m^* (modulo l'incertitude reliée au temps de vie de la particule de masse m^*). Par contre, dans un canal de réaction différent, les particules 4 et 5 pourraient être produites directement (sans passer par la résonance m^*) et alors la section efficace ne serait pas piquée à $M = m^*$, mais affecterait une forme étendue (appelée *fond d'espace de phase*). En réalité, on assiste à une superposition de ces deux processus.

Prob. 1.10

G Problèmes

Problème 1.1 Unités naturelles et analyse dimensionnelle

A L'annihilation d'un électron (e^-) et d'un positron (e^+) peut donner naissance à une paire quark-antiquark (q et \bar{q}). La section efficace de ce processus est donnée, dans le système d'unités naturelles et dans la limite ultrarelativiste, par la formule $\sigma = 4\pi\alpha^2 Q^2/M^2$, où (i) Q est la charge du quark ($\frac{1}{3}$ ou $\frac{2}{3}$) (ii) α est la constante de structure fine et (iii) M est la masse invariante de la paire e^+e^- . Restaurez dans cette formule les facteurs de c et de \hbar .

B Démontrez, par analyse dimensionnelle, que l'expression $e^2/\hbar 4\pi c$ est sans dimensions. Partez du fait que l'énergie potentielle électrostatique entre deux particules de charge e est $e^2/4\pi r$.

C Le positronium est un atome instable formé d'un électron et d'un positron. Il se désintègre rapidement en deux photons. Dans son état singulet (c'est-à-dire quand le spin total des deux particules est nul), sa vie moyenne inverse, calculée en électrodynamique quantique, est donnée par la formule $\tau^{-1} = \frac{1}{2}m\alpha^5$, où m est la masse de l'électron et α est la constante de structure fine. (i) Exprimez τ en restaurant les constantes \hbar et c et (ii) évaluez sa valeur numérique en secondes.

Problème 1.2 Désintégration du pion

Le méson π^- (masse 139,569 MeV) se désintègre par interaction faible en un muon μ^- (masse 105,659 MeV) et un antineutrino $\bar{\nu}_\mu$ (masse nulle). En supposant que le pion est initialement au repos, calculez la vitesse du muon produit par sa désintégration. Indice : la solution à ce problème est particulièrement simple dans le langage des quadrivecteurs.

Problème 1.3 Énergie de seuil d'une réaction

Une particule A, de masse m_A et d'énergie E , est incidente sur une particule B au repos, de masse m_B . La collision est inélastique et produit N particules dans l'état final. On définit l'énergie de *seuil* comme la valeur minimale de E en deçà de laquelle la réaction est impossible. Exprimez l'énergie de seuil en fonction seulement de m_A , de m_B et de la masse totale M des produits de la réaction. Notez que ce calcul est particulièrement simple dans le langage des quadrivecteurs, en utilisant la notion de masse invariante, surtout si on se place dans le référentiel du centre d'impulsion dans l'analyse des produits.

Appliquez ensuite la formule obtenue pour évaluer numériquement l'énergie de seuil (en MeV) de la réaction

$$p + p \rightarrow p + p + \pi^0 \quad (1.143)$$

où p désigne le proton (masse 938,280 MeV) et π^0 le pion neutre (masse 134,964 MeV).

Problème 1.4 Cinématique des collisions

Considérons un processus de collision élastique $1 + 2 \rightarrow 3 + 4$ où $m_1 = m_3$ et $m_2 = m_4$. Nous nous plaçons dans le référentiel du laboratoire, où la particule 2 est initialement au repos. La particule 1 est déviée d'un angle θ (l'angle de diffusion) par rapport à sa direction initiale.

A Déterminez l'énergie E_3 du projectile après la collision en fonction de l'énergie incidente E_1 et de l'angle θ . Indice : Appliquez la conservation de l'énergie et de la quantité de mouvement et calculez l'invariant $p_4^2 = (p_1 + p_2 - p_3)^2$. Vous pouvez utiliser un logiciel de calcul symbolique (ex. Mathematica) pour vous aider, mais le problème se réduit à la solution d'une équation quadratique.

B Simplifiez ce résultat dans le cas particulier $m_1 = 0$, correspondant à l'effet Compton (photon incident sur un électron). Partez de l'équation algébrique à résoudre et non de la solution à celle-ci.

Problème 1.5 Variables de Mandelstam

Considérons un processus de diffusion $1 + 2 \rightarrow 3 + 4$ où les réactants et les produits sont étiquetés de 1 à 4 comme indiqué. On définit les variables de Mandelstam s , t et u en fonction des quadri-vecteurs impulsion des particules impliquées :

$$s = (p_1 + p_2)^2 \quad t = (p_1 - p_3)^2 \quad u = (p_1 - p_4)^2$$

On suppose que l'énergie et la quantité de mouvement sont conservées ($p_1 + p_2 = p_3 + p_4$). Dans un processus de collision élastique, on suppose en outre, par convention, que les particules 1 et 3 sont du même type.

A Démontrez que les trois variables de Mandelstam ne sont pas indépendantes, mais respectent la contrainte

$$s + t + u = m_1^2 + m_2^2 + m_3^2 + m_4^2$$

B Lors d'une collision élastique, exprimez t en fonction du transfert de quantité de mouvement \mathbf{q} de la particule 1 à la particule 3 dans le repère du centre d'impulsion (ou centre de masse).

C Comme la section efficace est un invariant de Lorentz (tout comme les variables de Mandelstam), la quantité $d\sigma/dt$ est aussi un invariant de Lorentz. Montrez que, lors d'une collision élastique entre deux particules de même masse m , cette quantité est reliée à la section différentielle définie dans le repère du centre de masse par

$$\frac{d\sigma}{d\Omega} = \frac{s - 4m^2}{4\pi} \left| \frac{d\sigma}{dt} \right|$$

Problème 1.6 Section différentielle dans le référentiel du centre d'impulsion

La forme générale de la section différentielle est

$$d\sigma = \frac{|\mathcal{M}_{fi}|^2}{4\sqrt{(p_1 \cdot p_2)^2 - (m_1 m_2)^2}} \frac{d^3 p_3}{(2\pi)^3 2E_3} \cdots \frac{d^3 p_N}{(2\pi)^3 2E_N} (2\pi)^4 \delta^4(p_1 + p_2 - p_3 - \cdots p_N) \quad (1.144)$$

Considérons maintenant une collision à deux réactants et deux produits. Plaçons-nous dans le référentiel du centre d'impulsion des réactants, de sorte que $\mathbf{p}_1 + \mathbf{p}_2 = 0$.

A Démontrez que, dans ce référentiel,

$$\sqrt{(p_1 \cdot p_2)^2 - (m_1 m_2)^2} = |\mathbf{p}_1|(E_1 + E_2) \quad (1.145)$$

B Démontrez que la section différentielle dans ce référentiel est

$$\frac{d\sigma}{d\Omega} = \frac{|\mathcal{M}_{fi}|^2}{(8\pi)^2 E^2} \frac{E^2 - m_4^2 - m_3^2}{E^2 - m_1^2 - m_2^2} \quad (1.146)$$

où E l'énergie totale des réactants.

Problème 1.7 Section différentielle dans le référentiel du laboratoire

Considérons une collision $1 + 2 \rightarrow 3 + 4$, dans le référentiel du laboratoire (particule 2 au repos) et supposons que les produits sont les mêmes que les réactants ($m_3 = m_1$ et $m_4 = m_2$).

A En partant de la relation (1.139), démontrez que la section différentielle de diffusion (par unité d'angle solide de la particule 3, c'est-à-dire du projectile diffusé) est

$$\frac{d\sigma}{d\Omega} = \frac{1}{(8\pi)^2} \frac{\mathbf{p}_3^2 |\mathcal{M}_{fi}|^2}{m_2 |\mathbf{p}_1| (|\mathbf{p}_3|(E_1 + m_2) - |\mathbf{p}_1| E_3 \cos \theta)} \quad (1.147)$$

où E_i désigne l'énergie de la particule i . Notez qu'un changement de variable semblable à celui effectué en (1.80) est requis.

B Montrez que, dans le cas d'un projectile de masse nulle ($m_1 = 0$), comme dans l'effet Compton, on a la relation

$$\frac{1}{E_3} = \frac{1}{E_1} + \frac{1}{m_2} (1 - \cos \theta) \quad (1.148)$$

et que la formule ci-haut se réduit à

$$\frac{d\sigma}{d\Omega} = |\mathcal{M}_{fi}|^2 \left(\frac{E_3}{8\pi E_1 m_2} \right)^2 \quad (1.149)$$

Problème 1.8 Diffusion d'un neutron de basse énergie

Un neutron de basse énergie (1 keV) est incident sur un noyau de rayon $a = 5$ fm. Expliquez pourquoi on doit s'attendre, dans ce cas, à ce que la section différentielle de diffusion soit pratiquement isotrope, c'est-à-dire indépendante de l'angle de diffusion θ .

Problème 1.9 Droites de Kurie

Dans ce problème nous allons étudier la cinématique de l'émission β par un noyau. On considère un noyau instable (masse M) qui se désintègre en émettant un électron (masse m_e), un neutrino (masse m_ν , qu'on suppose non nulle dans les calculs) et un noyau de masse M' . Les masses des noyaux sont très grandes par rapport à m_e ou m_ν . Une énergie E est libérée par la désintégration, c'est-à-dire que

$$E = (M - M') \quad (1.150)$$

Cette énergie sert à créer l'électron, le neutrino et à donner de l'énergie cinétique aux trois produits (électron, neutrino et noyau). On suppose que $E \ll M$ et que les noyaux, en tout temps, demeurent non relativistes.

Une théorie de l'émission β a été proposée en 1934 par E. Fermi.^a Une modification de cette théorie fut ensuite proposée par Feynman et Gell-Mann dans les années 1950. Le cadre de cette théorie permet de calculer l'amplitude invariante suivante pour le processus :

$$|\mathcal{M}|^2 = \mathcal{G}^2 (p_1 \cdot p_2)(p_3 \cdot p_4) \quad (1.151)$$

où p_i est la quadri-impulsion de la particule i , m_i sa masse, et \mathcal{G} est une constante ayant les unités d'une surface. les particules 1 et 3 représentent les deux noyaux (avant et après la désintégration) alors que les particules 2 et 4 sont l'électron et le neutrino.

A Expliquer pourquoi on peut négliger l'énergie cinétique du noyau produit, par rapport à celle donnée au neutrino et à l'électron. Autrement dit, l'électron et le neutrino se partagent l'énergie E .

B Pour une énergie libérée E fixe, quel est l'intervalle de valeurs possibles pour l'énergie de l'électron?

C La quantité de mouvement de l'électron émis n'est pas fixe, mais on peut définir un taux d'émission par intervalle de quantité de mouvement de l'électron. Montrez que le taux d'émission par unité de quantité de mouvement de l'électron est

$$\frac{d\omega}{d|\mathbf{p}_e|} = \frac{\mathcal{G}^2}{32\pi^3} \mathbf{p}_e^2 (E - E_e) \sqrt{(E - E_e)^2 - m_\nu^2} \quad (1.152)$$

Vous devez pour cela faire l'approximation que les masses des noyaux (M et M') sont beaucoup plus grandes que l'énergie libérée.

D Il découle de la partie précédente qu'un graphique de $\sqrt{d\omega/d|\mathbf{p}_e|}/|\mathbf{p}_e|$ en fonction de E_e est une droite (dite *droite de Kurie*^b) si la masse du neutrino est nulle. Faites un tel graphique (par

exemple, à l'aide de Mathematica) pour les paramètres suivants : $E = 1,1 m_e$ et $m_\nu/m_e = 0,01$ et $0,05$. Ne vous préoccupez pas du préfacteur $\mathcal{G}^2/32\pi^3$. L'analyse des droites de Kurie obtenues expérimentalement a permis d'obtenir une limite supérieure de $2,2$ eV à la masse des neutrinos. L'expérience KATRIN, en Allemagne, espère abaisser cette limite d'un ordre de grandeur.

- a. E. Fermi, Zeitschrift für Physik, **88**, 161 (1934). Une traduction anglaise est disponible : American Journal of Physics **36**, 1150 (1968).
 b. Oui, c'est un 'K'. Rien à voir avec Curie.

Problème 1.10 Espace des phases

Considérons une collision de deux particules (1 et 2) qui mène à un état final formé de trois particules (numérotées 3, 4 et 5). Le problème est de déterminer la distribution de la masse invariante $m = \sqrt{(p_4 + p_5)^2}$ des deux dernières particules. Autrement dit, quelle est la densité d'états finaux $\rho(m)$ pour lesquels la masse invariante M_{45} est égale à m ? La réaction $p\pi^- \rightarrow n\pi^+\pi^-$ est un exemple de ce processus. Dans ce cas, les masses sont (en MeV) $m_1 = 938$, $m_3 = 939$ et $m_2 = m_4 = m_5 = 140$. Considérez cependant que les masses m_i des particules sont des paramètres généraux, comme la masse invariante $s = (p_1 + p_2)^2$ des réactants.

La distribution $\rho(m)$ s'obtient en faisant la somme sur tous les états finaux possibles, avec la contrainte que l'énergie et la quantité de mouvement sont conservées et que la masse invariante des particules 4 et 5 vaut m . D'où l'expression

$$\rho(m) = \int \frac{d^3 p_3}{(2\pi)^3 2E_3} \frac{d^3 p_4}{(2\pi)^3 2E_4} \frac{d^3 p_5}{(2\pi)^3 2E_5} (2\pi)^4 \delta^4(p_1 + p_2 - p_3 - p_4 - p_5) \delta(m - \sqrt{(p_4 + p_5)^2}) \quad (1.153)$$

A Pour des valeurs données de s et des masses, quelle est la valeur maximale que m peut admettre? Justifiez votre réponse sans calcul complexe.

B Quelle est sa valeur minimale? Justifiez votre réponse sans calcul complexe.

C Démontrez que

$$\rho(m) = \frac{1}{32\pi^3} \frac{\sqrt{[m^2 - (m_4 + m_5)^2][m^2 - (m_4 - m_5)^2]}}{m\sqrt{s}} \sqrt{\left(\frac{s + m_3^2 - m^2}{2\sqrt{s}}\right)^2 - m_3^2} \quad (1.154)$$

Indice : commencez par évaluer l'intégrale sur \mathbf{p}_4 et \mathbf{p}_5 , pour une valeur fixe de \mathbf{p}_3 , et pour ce faire placez-vous dans le référentiel où $\mathbf{p}_1 + \mathbf{p}_2 - \mathbf{p}_3 = 0$. Une fois cette partie de l'intégrale effectuée, il faut exprimer le résultat de manière invariante de Lorentz en fonction des variables restantes (\mathbf{p}_3) et procéder à l'intégrale sur \mathbf{p}_3 . Cette dernière intégrale se fait facilement dans le référentiel où $\mathbf{p}_1 + \mathbf{p}_2 = 0$. Aucune évaluation de primitive (c.-à-d. d'intégrale indéfinie) n'est requise dans ce calcul. Tout se fait par le jeu des fonctions delta et l'invariance par rotation dans des référentiels bien choisis.

D Portez cette fonction $\rho(m)$ en graphique (à l'aide de Mathematica, par exemple) en utilisant les masses ci-haut pour la réaction $p\pi^- \rightarrow n\pi^+\pi^-$ et une valeur d'énergie initiale de $E = 1\,950$ MeV dans le référentiel du centre de masse des particules 1 et 2.

CHAPITRE 2

CHAMPS QUANTIQUES ET INTERACTIONS

La théorie quantique des champs est l'outil de base du modèle standard. Ce chapitre se veut une introduction à cet outil, via la théorie du champ scalaire, décrivant des bosons de spin 0, et celle du champ de Schrödinger, décrivant des fermions non relativistes. Malgré leur relative simplicité, ces théories permettent d'introduire tous les concepts de base, y compris les diagrammes de Feynman. Cependant, il faut garder à l'esprit que les particules élémentaires du tableau de la page 14 sont décrites par des théories légèrement différentes : le champ de Dirac (voir chapitre 3) pour les fermions, et les champs de jauge pour les bosons (voir les sections 4.B et 5.D). Seul le champ de Higgs est décrit par un champ scalaire.

A

Théorie du champ scalaire

2.A.1 Un modèle simple en dimension 1

Comme prélude à l'étude des champs quantiques, nous allons étudier un modèle simple décrivant les vibrations d'un cristal en dimension 1. Ce modèle simple servira de base microscopique à la théorie du champ scalaire. L'objectif n'est pas du tout de décrire de manière réaliste les phonons existant dans les solides, qui sont caractérisés par plusieurs bandes et polarisations, mais uniquement d'illustrer comment des excitations, ou particules peuvent émerger d'un modèle comportant un nombre quasi infini de degrés de liberté.

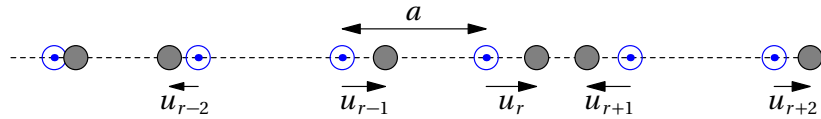
Considérons une chaîne de N billes (ou «atomes») de masse μ distribuées le long d'une droite en dimension 1 (Fig. 2.1). Ces billes sont reliées par des ressorts de fréquence caractéristique Ω à des points d'ancrage régulièrement espacés, formant un réseau unidimensionnel de pas a . En plus, les billes sont reliées à leurs voisins immédiats par d'autres ressorts, de fréquence caractéristique Γ . À l'équilibre, les billes demeurent sur leurs points d'ancrage. Mais un déplacement quelconque par rapport à cet équilibre cause l'apparition d'ondes longitudinales se propageant le long de cette

chaîne. Si u_r désigne le déplacement de la bille n° r par rapport à sa position d'équilibre, alors le lagrangien de ce système est

$$L = \frac{1}{2}\mu \sum_{r=1}^N \{ \dot{u}_r^2 - \Omega^2 u_r^2 - \Gamma^2 (u_r - u_{r+1})^2 \} \quad (2.1)$$

Nous pouvons supposer, pour simplifier, que ce système possède des conditions aux limites périodiques, c'est-à-dire que $u_{N+1} \equiv u_1$. La longueur du système est $\ell = Na$.

FIGURE 2.1



Les équations de Lagrange associées sont les suivantes :

$$\frac{d}{dt} \frac{\partial L}{\partial \dot{u}_r} - \frac{\partial L}{\partial u_r} = 0 = \ddot{u}_r + \Omega^2 u_r + \Gamma^2 (2u_r - u_{r+1} - u_{r-1}) \quad (2.2)$$

Il ne faut pas oublier, en dérivant ces équations, que le terme impliquant u_s dans le dernier terme du lagrangien (2.1) apparaît dans deux des termes de la somme : $r = s$ et $r = s - 1$. Ces équations forment un système d'équations linéaires couplées. Cherchons-en une solution ayant la forme d'une onde progressive, c'est-à-dire

$$u_r = A e^{i(qr - \omega t)} \quad (2.3)$$

Comme la variable u_r est réelle, il est sous-entendu ici que nous prenons la partie réelle de cette expression et que l'amplitude A peut être complexe. Le paramètre q ne peut prendre que des valeurs comprises dans l'ensemble $2\pi\mathbb{Z}/N$, en raison de la condition de périodicité. De plus, deux valeurs de q séparées de 2π sont entièrement équivalentes. En substituant cette forme dans l'équation (2.2), on trouve la condition

$$-\omega^2 + \Omega^2 + \Gamma^2 [2 - e^{iq} - e^{-iq}] = 0 \quad (2.4)$$

La fréquence de l'onde est donc donnée par l'expression suivante :

$$\omega_q = \sqrt{\Omega^2 + 2\Gamma^2(1 - \cos q)} = \sqrt{\Omega^2 + 4\Gamma^2 \sin^2(q/2)} \quad (2.5)$$

Cette dernière relation constitue la *relation de dispersion*, soit la relation entre la fréquence et le nombre d'onde d'une onde progressive.

Comme l'équation (2.2) est linéaire, sa solution générale est une combinaison linéaire des solutions particulières trouvées ci-dessus, pour toutes les valeurs distinctes de q :

$$u_r(t) = \frac{1}{\sqrt{N}} \sum_q \tilde{u}_q e^{i(qr - \omega_q t)} \quad (2.6)$$

où \tilde{u}_q est une amplitude qui doit être déterminée d'après les conditions initiales du problème. Le facteur $1/\sqrt{N}$ a été introduit par convention, afin de fixer la normalisation de chacune des solutions indépendantes qui sont combinées. Cette solution générale est une superposition de modes de fréquences définies ω_q . Du point de vue classique, le problème est dès lors résolu.

Les ondes qui se propagent ainsi sont des *modes collectifs* et ont, en mécanique quantique, un caractère corpusculaire autant qu'ondulatoire. Autrement dit, l'énergie contenue dans ces vibrations

est quantifiée en multiples entiers de $\hbar\omega_q$, et chaque quantum d'oscillation se comporte comme une particule. En physique de solide, les quanta de vibration du réseau cristallin sont appelés *phonons*, et c'est ainsi que nous pourrions appeler les particules qui émergent du modèle simple (2.1), même si à proprement parler on ne peut pas parler de phonons ici en raison du terme en Ω^2 qui n'existe pas sur un véritable réseau cristallin. Mais ces phonons ne sont pas les billes d'origine qui, elles, sont liées et ne peuvent se propager; ce sont plutôt des particules *émergentes*.

2.A.2 Limite continue

Supposons que le pas de réseau a soit très petit par rapport à nos moyens d'observation, de sorte qu'on puisse traiter approximativement le système de billes comme un continuum. Nous allons remplacer la variable u_r par une fonction $u(x)$, où x est une variable continue dans l'intervalle $[0, L]$ qui prend les valeurs $x = ra$ dans le cas discret ($Na = \ell$ et $r = 1, 2, \dots, N$). La différence finie $u_{r+1} - u_r$ sera alors remplacée par une dérivée : $a\partial_x u$. Le lagrangien (2.1) peut alors être remplacé par l'expression suivante :

$$L = \frac{1}{2}\mu \int_0^\ell dx \frac{1}{a} [\dot{u}^2 - \Omega^2 u^2 - \Gamma^2 a^2 (\partial_x u)^2] \quad (2.7)$$

Il est alors pratique de définir un champ $\phi(x) = u(x)\sqrt{\mu/a}$, de sorte que

$$L = \frac{1}{2} \int_0^\ell dx [\dot{\phi}^2 - \Omega^2 \phi^2 - c^2 (\partial_x \phi)^2] \quad (2.8)$$

où nous avons également introduit la vitesse caractéristique $c = \Gamma a$. On définit habituellement une densité lagrangienne \mathcal{L} telle que

$$L = \int dx \mathcal{L} \quad \mathcal{L} = \dot{\phi}^2 - \Omega^2 \phi^2 - c^2 (\partial_x \phi)^2 \quad (2.9)$$

Équation de Lagrange On peut également appliquer la limite continue directement aux équations de Lagrange (2.2) :

$$\frac{\partial^2 u}{\partial t^2} + \Omega^2 u - a^2 \Gamma^2 \frac{\partial^2 u}{\partial x^2} = 0 \quad (2.10)$$

où nous avons utilisé la représentation discrète de la dérivée seconde (à l'ordre le plus bas en a) :

$$\frac{\partial^2 u}{\partial x^2} \approx \frac{u_{r+1} + u_{r-1} - 2u_r}{a^2} \quad (2.11)$$

En fonction du champ ϕ , l'équation de Lagrange demeure la même :

$$\frac{\partial^2 \phi}{\partial t^2} + \Omega^2 \phi - c^2 \frac{\partial^2 \phi}{\partial x^2} = 0 \quad (2.12)$$

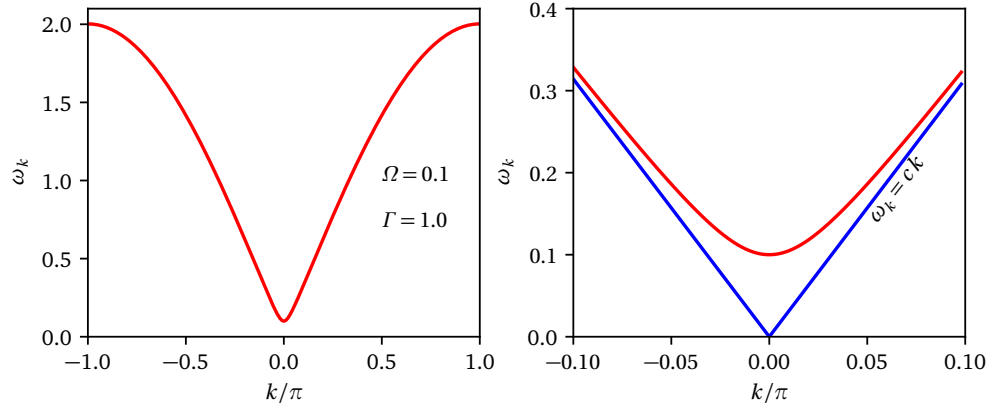


FIGURE 2.2

En haut : relation de dispersion (2.5). En bas, élargissement de la partie basse énergie ; les droites correspondent à la dispersion linéaire de particules sans masse.

Relation de dispersion Prenons maintenant la limite continue de la relation de dispersion (2.5). Nous devons pour cela supposer que le nombre d'onde q est très petit par rapport à π . On trouve alors

$$\omega_q = \sqrt{\Omega^2 + 4\Gamma^2 \sin^2(q/2)} \approx \sqrt{\Omega^2 + \Gamma^2 q^2} = \sqrt{m^2 c^4 + c^2 p^2} \quad p := \frac{q}{a}, \quad m := \frac{\Omega}{c} \quad (2.13)$$

On a introduit ici un véritable vecteur d'onde p et la masse m , ayant les unités voulues. L'approximation du continuum revient à demander que p soit petit en comparaison de a^{-1} , le pas de réseau inverse. On trouve en tout cas l'expression relativiste de l'énergie $E(p)$ d'une particule d'impulsion p et de masse m . La relation de dispersion du système discret, et sa limite de faible vecteur d'onde sont illustrées à la figure 2.2. Les ondes qui se propagent dans ce modèle peuvent donc être assimilées à des particules de masse m . Cette interprétation sera clarifiée plus bas.

Hamiltonien Dans le système discret d'origine, la relation entre le lagrangien et le hamiltonien est la suivante :

$$H = \sum_r p_r \dot{u}_r - L \quad \text{où} \quad p_r = \frac{\partial L}{\partial \dot{u}_r} \quad (2.14)$$

Comment ce passage est-il réalisé dans la limite continue ? Dans le cas quasi continu, mais encore discret, on pourrait généralement écrire

$$L = \sum_r a \mathcal{L}(\phi(x_r), \dot{\phi}(x_r)) \quad (2.15)$$

où les différentes valeurs $x_r = r a$ de la coordonnée spatiale sont espacées régulièrement. Le moment conjugué à la variable $\phi(x_r)$ est alors

$$\frac{\partial L}{\partial \dot{\phi}(x_r)} = a \pi(x_r) \quad \text{où} \quad \pi(x) := \frac{\partial \mathcal{L}}{\partial \dot{\phi}(x)} \quad (2.16)$$

et le crochet de Poisson associé est

$$[\phi(x_r), \pi(x_s)]_P = \frac{1}{a} \delta_{rs} \quad (2.17)$$

Dans la limite continue, l'expression δ_{rs}/a devient équivalente à une fonction delta de Dirac :¹

$$\frac{1}{a} \delta_{rs} \rightarrow \delta(x_r - x_s) \quad (2.20)$$

Donc, le crochet de Poisson prend la forme suivante pour un système continu :

$$[\phi(x), \pi(x')]_P = \delta(x - x') \quad (2.21)$$

Le hamiltonien prend alors la forme suivante :

$$H = \sum_r a \pi(x_r) \dot{\phi}(x_r) - L = \int dx [\pi(x) \dot{\phi}(x) - \mathcal{L}] \quad (2.22)$$

On peut donc le représenter comme l'intégrale d'une densité hamiltonienne \mathcal{H} :

$$H = \int dx \mathcal{H} \quad \text{où} \quad \boxed{\mathcal{H} = \pi(x) \dot{\phi}(x) - \mathcal{L}} \quad (2.23)$$

Généralisation à trois dimensions Le modèle ci-dessus peut se généraliser à trois dimensions d'espace en supposant que le champ ϕ dépend maintenant de la position \mathbf{r} et que le lagrangien associé prend la forme suivante :

$$L = \frac{1}{2} \int d^3r \{ \dot{\phi}^2 - \Omega^2 \phi^2 - c^2 (\nabla \phi)^2 \} \quad (2.24)$$

L'équation de Lagrange devient alors

$$\frac{\partial^2 \phi}{\partial t^2} + \Omega^2 \phi - c^2 \nabla^2 \phi = 0 \quad (2.25)$$

Notons que dans cette généralisation à trois dimensions, on perd quelque peu l'interprétation originale : le champ ϕ peut toujours être interprété comme une mise à l'échelle du déplacement des billes, mais dans une direction seulement. Un modèle de vibrations plus réaliste permettrait aux billes de se déplacer dans toutes les directions, ce qui n'est pas le cas ici. Gardons à l'esprit que le modèle des billes n'est utile ici que pour motiver l'apparition de modes collectifs de vibration.

1. On le voit facilement en constatant que

$$\sum_s a \left(\frac{1}{a} \delta_{rs} \right) f(x_s) = f(x_r) \quad (2.18)$$

et que cette expression prend également la forme suivante dans la limite continue :

$$\int dx_s \delta(x_r - x'_s) f(x'_s) = f(x_r) \quad (2.19)$$

En trois dimensions, le passage vers le hamiltonien se fait de la manière suivante :

$$H = \int d^3x \mathcal{H} \quad \boxed{\mathcal{H} = \pi(\mathbf{r})\dot{\phi}(\mathbf{r}) - \mathcal{L}} \quad \text{où} \quad \boxed{\pi(\mathbf{r}) = \frac{\partial \mathcal{L}}{\partial \dot{\phi}(\mathbf{r})}} \quad (2.26)$$

De plus, on a le crochet de Poisson suivant :

$$\boxed{[\phi(\mathbf{r}), \pi(\mathbf{r}')]_P = \delta(\mathbf{r} - \mathbf{r}')} \quad (2.27)$$

2.A.3 Équation de Klein-Gordon

L'équation (2.25) est une modification de l'équation d'onde habituelle et porte le nom d'équation de *Klein-Gordon*. Elle a été proposée la première fois dans le contexte de la recherche d'une équation d'onde relativiste pour les ondes de matière (voir ci-dessous). On peut récrire cette équation de la manière suivante, en définissant le paramètre $m = \Omega/c^2$:

$$\frac{1}{c^2} \frac{\partial^2 \phi}{\partial t^2} + m^2 c^2 \phi - \nabla^2 \phi = 0 \quad m := \frac{\Omega}{c^2} \quad (2.28)$$

Une propriété remarquable de cette équation est qu'elle respecte l'invariance de Lorentz, si la vitesse caractéristique c coïncide avec la vitesse limite (vitesse de la lumière). En effet, l'équation prend la forme suivante dans le langage des quadrivecteurs (on pose maintenant $c = 1$) :

$$\boxed{\partial_\mu \partial^\mu \phi + m^2 \phi = 0} \quad (2.29)$$

En supposant que cette équation admette des solutions de type onde plane, qui ont la forme $\phi = \phi_0 e^{-ip_\mu x^\mu}$ (ϕ_0 étant une constante), on trouve, en substituant dans l'équation, la condition suivante sur le quadrivecteur d'onde (ou quadri-impulsion) :

$$(-p_\mu p^\mu + m^2)\phi_0 = 0 \quad \text{ou encore} \quad p^2 = E^2 - \mathbf{p}^2 = m^2 \quad (2.30)$$

Autrement dit, l'énergie-impulsion p^μ de l'onde obéit à la relation de dispersion de particules relativistes de masse m (d'où la notation choisie pour m). La masse m n'a bien sûr rien à voir avec la masse μ des billes du modèle d'origine. Elle est essentiellement reliée à la fréquence caractéristique du ressort qui maintient les billes autour de leur point d'ancrage. Autrement dit : dans le cas où toutes les billes sont déplacées en phase et d'une même distance par rapport à leur position d'équilibre, Ω est la fréquence d'oscillation du système, et un quantum de cette oscillation aurait une énergie $\hbar\Omega$. Ce quantum est associé à une particule au repos dont l'énergie de masse est $mc^2 = \hbar\Omega$ (nous allons dorénavant poser $\hbar = 1$, comme dans le reste de ce recueil).

Nous avons obtenu ici un résultat profond : nous sommes partis d'un système discret (le réseau de billes), qui ne possède pas d'invariance de Lorentz. En effet, les billes constituent une sorte d'«éther», de milieu de propagation, et leur référentiel au repos est privilégié. En prenant la limite

continue, nous avons obtenu une théorie effective qui, elle, possède l'invariance de Lorentz. Autrement dit, l'une des symétries les plus fondamentales de la physique apparaît comme un aspect émergent, approximatif, d'une théorie discrète plus fondamentale définie à une échelle de longueur trop petite pour être accessible. Il ne faut bien sûr pas conclure de cet exercice que l'univers est vraiment fait de billes microscopiques! Mais il faut garder à l'esprit qu'une théorie impliquant des champs continus n'est vraisemblablement qu'une approximation d'une théorie encore plus fondamentale, qui nous échappe pour le moment.

Motivation originale pour l'équation de Klein-Gordon

L'équation de Klein-Gordon fut proposée en 1926 pour décrire les ondes de matière relativistes par plusieurs physiciens en 1926, notamment Klein² et Gordon³ L'équation de Schrödinger habituelle

pour les particules libres

$$i \frac{\partial \psi}{\partial t} = -\frac{1}{2m} \nabla^2 \psi \quad (2.31)$$

n'est pas compatible avec la théorie de la relativité restreinte. Elle décrit des particules dont l'énergie et la quantité de mouvement sont reliées par $E = \mathbf{p}^2/2m$. Cela se constate immédiatement par la correspondance entre énergie, quantité de mouvement et les opérateurs différentiels, soit

$$E \rightarrow i \frac{\partial}{\partial t} \quad \mathbf{p} \rightarrow -i \nabla \quad \text{ou encore} \quad p_\mu \rightarrow i \partial_\mu \quad (2.32)$$

En suivant de près les idées originales de Louis de Broglie et en appliquant la relation relativiste entre énergie et quantité de mouvement, soit $p_\mu p^\mu = m^2$, à la correspondance ci-dessus, on obtient bel et bien l'équation de Klein-Gordon (2.29). Cette équation est très semblable à l'équation d'onde ordinaire, sauf pour un terme supplémentaire impliquant la masse. Cependant, l'équation de Klein-Gordon ne peut pas correctement décrire les électrons relativistes. Non seulement ne contient-elle pas le spin, mais en la couplant on champ électromagnétique, on ne trouve pas la forme correcte de la structure fine du spectre de l'atome d'hydrogène.

Absence de courant de probabilité

Un paradoxe posé par l'équation de Klein-Gordon est que, étant du deuxième ordre dans le temps, l'énergie de la particule peut être à la fois négative ou positive : $E = \pm \sqrt{\mathbf{p}^2 + m^2}$. On ne savait pas, à l'époque, comment interpréter ces solutions à énergie négative. Reliée à cela est l'impossibilité d'interpréter le module carré de ψ comme une densité de probabilité. En effet, si ψ obéit à l'équation de Schrödinger, on peut définir une densité positive $\rho = |\psi|^2$ et une densité de courant $\mathbf{J} = -(i/2m)(\psi^* \nabla \psi - \nabla \psi^* \psi)$ qui respectent l'équation de continuité :

$$\frac{\partial \rho}{\partial t} + \nabla \cdot \mathbf{J} = 0 \quad (2.33)$$

En effet, en substituant les définitions de ρ et \mathbf{J} dans cette dernière équation, on trouve

$$\begin{aligned} \frac{\partial \rho}{\partial t} + \nabla \cdot \mathbf{J} &= \dot{\psi}^* \psi + \psi^* \dot{\psi} - \frac{i}{2m} \psi^* \nabla^2 \psi + \frac{i}{2m} \nabla^2 \psi^* \psi \\ &= [\dot{\psi} - i \frac{1}{2m} \nabla^2 \psi]^* \psi + \psi^* [\dot{\psi} - i \frac{1}{2m} \nabla^2 \psi] = 0 \end{aligned} \quad (2.34)$$

2. O. Klein, *Z. f. Phys.* **37**, 895 (1926).

3. W. Gordon, *Z. f. Phys.* **40**, 117 (1926).

où l'annulation provient de l'équation de Schrödinger. Dans le cas de l'équation de Klein-Gordon, l'invariance relativiste de l'équation nous permet de définir une densité et un courant qui forment un quadrivecteur :

$$J^\mu = (\rho, \mathbf{J}) = i(\psi^* \partial^\mu \psi - \partial^\mu \psi^* \psi) \quad (2.35)$$

(le facteur i est nécessaire afin que la densité de courant soit réelle). L'équation de continuité est alors valable en vertu de l'équation de Klein-Gordon :

$$\begin{aligned} \partial_\mu J^\mu &= i \partial_\mu (\psi^* \partial^\mu \psi - \partial^\mu \psi^* \psi) \\ &= i (\psi^* \partial_\mu \partial^\mu \psi - \partial_\mu \partial^\mu \psi^* \psi) \\ &= i (-m^2 \psi^* \psi + m^2 \psi^* \psi) = 0 \end{aligned} \quad (2.36)$$

La densité ρ est donc

$$\rho = i(\psi^* \dot{\psi} - \dot{\psi}^* \psi) \quad (2.37)$$

mais cette quantité n'est pas définie positive et ne peut donc pas être interprétée comme une densité de probabilité.

En fait, la raison fondamentale pour laquelle cette équation ne peut s'appliquer aux électrons est un peu plus subtile : cette équation, pour être fondamentale, ne peut décrire que des bosons de spin zéro. La description relativiste des particules de spin 1/2 est plutôt fournie par l'équation de Dirac.

2.A.4 Invariance de Lorentz de la théorie du champ scalaire

Retournons à l'action (2.24), cette fois en adoptant un langage covariant en dimension 3+1 et en posant $c = 1$:

$$S = \frac{1}{2} \int d^4x (\partial_\mu \phi \partial^\mu \phi - m^2 \phi^2) \quad (2.38)$$

On constate que cette expression est invariante lors des transformations de Lorentz, pourvu que le champ ϕ lui-même soit un invariant. Que veut-on dire par cette dernière condition? Le fait que le champ ϕ soit un invariant de Lorentz signifie que $\phi(x)$, considéré comme fonction de l'espace-temps, se transforme comme suit lorsqu'on procède à une transformation de Lorentz $x^\mu \rightarrow x'^\mu = \Lambda^\mu{}_\nu x^\nu$:

$$\phi(x) \rightarrow \phi'(x') = \phi(x)$$

Autrement dit, à un événement de l'espace-temps défini par la quadriposition x dans le référentiel S , le champ ϕ prend une valeur scalaire précise, qui ne dépend que de cet événement et non pas du référentiel utilisé pour le repérer dans l'espace-temps. En passant à un autre référentiel S' , la valeur ϕ' du champ au même événement est la même que dans S , c'est-à-dire que $\phi' = \phi$.

Le champ ϕ étant un invariant de Lorentz, il est alors clair que $\partial_\mu \phi$ est un quadrivecteur covariant et que les différents termes de l'action (2.38) sont invariants de Lorentz, incluant la mesure d'intégration d^4x sur l'espace-temps. Donc l'action S elle-même est un invariant de Lorentz. Comme les équations classiques du mouvement sont obtenues de la variation de l'action S , l'invariance de S entraîne nécessairement que les équations du mouvement ont la même forme dans tous les référentiels. C'est le cas de l'équation de Klein-Gordon (2.29).

Application covariante du principe variationnel Voyons comment l'équation de Klein-Gordon, dans sa forme covariante (2.29), dérive du principe variationnel. Il n'y a rien de nouveau ici, sauf dans la notation. En appliquant une variation $\delta\phi$ du champ, la variation de l'action est, au premier ordre en $\delta\phi$,

$$\delta S = \int d^4x (\partial_\mu \phi \partial^\mu \delta\phi - m^2 \phi \delta\phi) \quad (2.39)$$

En intégrant par parties le premier terme, on trouve

$$\delta S = \int d^4x (-\partial^\mu \partial_\mu \phi - m^2 \phi) \delta\phi + \int d^4x \partial_\mu (\partial^\mu \phi \delta\phi) \quad (2.40)$$

Le deuxième terme est l'intégrale sur l'espace-temps d'une quadridivergence. Le théorème de la divergence en dimension 4 s'applique et cette intégrale peut être ramenée à une intégrale sur une hypersurface dans l'espace-temps :

$$\int d^4x \partial_\mu (\partial^\mu \phi \delta\phi) = \oint da_\mu (\partial^\mu \phi \delta\phi) \quad (2.41)$$

Cette dernière intégrale pourrait prendre la forme d'une intégrale sur tout l'espace aux temps extrêmes t_1 et t_2 (comme ci-haut), plus une intégrale sur tous les temps, évaluée à la frontière spatiale du domaine considéré, à moins qu'on utilise des conditions aux limites périodiques qui font que le domaine spatial n'a pas de frontières. Dans tous les cas, on suppose que la variation $\delta\phi$ s'annule sur la frontière : cette condition fait partie du principe variationnel lui-même, qui est «local» dans l'espace-temps, en dépit des apparences. On peut par exemple supposer que le domaine d'intégration dans l'espace-temps est infini dans toutes les directions et que la variation $\delta\phi$ s'annule suffisamment rapidement à l'infini pour que le terme de surface puisse toujours être ignoré. Pour que la variation δS s'annule, quelle que soit la variation $\delta\phi$, il faut donc que l'expression entre parenthèses dans le premier terme soit nulle, ce qui mène précisément à l'équation de Klein-Gordon.

B Quantification du champ scalaire

Le passage de la mécanique classique à la mécanique quantique se fait en remplaçant les variables dynamiques qui figurent dans le hamiltonien par des opérateurs hermitiens. De plus, les crochets de Poisson deviennent des commutateurs (divisés par i) :

$$[A, B]_P \rightarrow \frac{1}{i\hbar} (\hat{A}\hat{B} - \hat{B}\hat{A}) \quad (2.42)$$

Par exemple, la relation de commutation canonique entre deux variables conjuguées q et p , soit $[q, p]_P = 1$, devient

$$\frac{1}{i} (\hat{q}\hat{p} - \hat{p}\hat{q}) = 1 \quad \text{ou encore} \quad [\hat{q}, \hat{p}] = i\hbar \quad (2.43)$$

Dans ce qui suit, nous adopterons les unités naturelles ($\hbar = 1$) et nous omettrons d'affubler les opérateurs d'un accent circonflexe, afin de ne pas alourdir la notation. La distinction entre variable classique et opérateur sera claire en fonction du contexte.

2.B.1 Hamiltonien

Le lagrangien du champ scalaire en trois dimensions est

$$L = \frac{1}{2} \int d^3 r [\dot{\phi}^2 - m^2 \phi^2 - (\nabla \phi)^2] \quad (2.44)$$

Le moment conjugué correspondant est $\pi(\mathbf{r}) = \dot{\phi}(\mathbf{r})$ et donc le hamiltonien du champ est

$$H = \int d^3 r \mathcal{H} \quad \mathcal{H} = \frac{1}{2} [\pi(\mathbf{r})^2 + m^2 \phi^2 + (\nabla \phi)^2] \quad (2.45)$$

Les quantités $\phi(\mathbf{r})$ et $\pi(\mathbf{r})$ sont maintenant des opérateurs, qui obéissent aux relations de commutation suivantes :

$$[\phi(\mathbf{r}), \pi(\mathbf{r}')] = i \delta(\mathbf{r} - \mathbf{r}') \quad [\phi(\mathbf{r}), \phi(\mathbf{r}')] = [\pi(\mathbf{r}), \pi(\mathbf{r}')] = 0 \quad (2.46)$$

Afin de simplifier la forme de ce hamiltonien, nous allons l'exprimer en fonction des transformées de Fourier

$$\phi(\mathbf{r}) = \frac{1}{\mathcal{V}} \sum_{\mathbf{p}} \phi_{\mathbf{p}} e^{i\mathbf{p}\cdot\mathbf{r}} \quad \phi_{\mathbf{p}} = \int d^3 r \phi(\mathbf{r}) e^{-i\mathbf{p}\cdot\mathbf{r}} \quad (2.47)$$

et

$$\pi(\mathbf{r}) = \frac{1}{\mathcal{V}} \sum_{\mathbf{p}} \pi_{\mathbf{p}} e^{i\mathbf{p}\cdot\mathbf{r}} \quad \pi_{\mathbf{p}} = \int d^3 r \pi(\mathbf{r}) e^{-i\mathbf{p}\cdot\mathbf{r}} \quad (2.48)$$

Notons que les opérateurs $\phi(\mathbf{r})$ et $\pi(\mathbf{r})$ sont hermitiens, ce qui n'est pas le cas de leurs transformées de Fourier :

$$\phi_{\mathbf{p}}^\dagger = \phi_{-\mathbf{p}} \quad \pi_{\mathbf{p}}^\dagger = \pi_{-\mathbf{p}} \quad (2.49)$$

En insérant ces transformées de Fourier dans l'expression du hamiltonien, on trouve

$$\begin{aligned} H &= \frac{1}{2\mathcal{V}^2} \sum_{\mathbf{p}, \mathbf{p}'} \int d^3 r e^{i\mathbf{r}\cdot(\mathbf{p}+\mathbf{p}')} [\pi_{\mathbf{p}'} \pi_{\mathbf{p}} + m^2 \phi_{\mathbf{p}'} \phi_{\mathbf{p}} - \mathbf{p}' \cdot \mathbf{p} \phi_{\mathbf{p}'} \phi_{\mathbf{p}}] \\ &= \frac{1}{2\mathcal{V}} \sum_{\mathbf{p}, \mathbf{p}'} \delta_{\mathbf{p}, -\mathbf{p}'} [\pi_{\mathbf{p}'} \pi_{\mathbf{p}} + m^2 \phi_{\mathbf{p}'} \phi_{\mathbf{p}} - \mathbf{p}' \cdot \mathbf{p} \phi_{\mathbf{p}'} \phi_{\mathbf{p}}] \\ &= \frac{1}{2\mathcal{V}} \sum_{\mathbf{p}} [\pi_{-\mathbf{p}} \pi_{\mathbf{p}} + m^2 \phi_{-\mathbf{p}} \phi_{\mathbf{p}} + \mathbf{p}^2 \phi_{-\mathbf{p}} \phi_{\mathbf{p}}] \\ &= \frac{1}{2\mathcal{V}} \sum_{\mathbf{p}} [\pi_{\mathbf{p}}^\dagger \pi_{\mathbf{p}} + \omega_{\mathbf{p}}^2 \phi_{\mathbf{p}}^\dagger \phi_{\mathbf{p}}] \quad \omega_{\mathbf{p}} := \sqrt{\mathbf{p}^2 + m^2} \end{aligned} \quad (2.50)$$

Cette dernière expression fait immédiatement penser au hamiltonien d'un ensemble d'oscillateurs harmoniques découplés (un oscillateur pour chaque valeur de \mathbf{p}) de fréquences $\omega_{\mathbf{p}}$ et de masse \mathcal{V} . La suite logique des choses est d'introduire les opérateurs d'échelles $a_{\mathbf{p}}$ et $a_{\mathbf{p}}^\dagger$ correspondants. Mais comme les transformées de Fourier à \mathbf{p} et à $-\mathbf{p}$ sont reliées par une conjugaison hermitienne, nous devons être prudents et s'assurer d'utiliser les bonnes relations de commutation.

Justement, les relations de commutation entre les transformées de Fourier sont les suivantes :

$$[\phi_{\mathbf{p}}, \pi_{\mathbf{p}'}] = \int d^3r d^3r' e^{-i(\mathbf{p}\cdot\mathbf{r} + \mathbf{p}'\cdot\mathbf{r}')} [\phi(\mathbf{r}), \pi(\mathbf{r}')] \quad (2.51)$$

$$= i \int d^3r e^{-i\mathbf{r}\cdot(\mathbf{p} + \mathbf{p}')} \quad (2.52)$$

$$= i \mathcal{V} \delta_{\mathbf{p}, -\mathbf{p}'} \quad (2.53)$$

ou encore, en remplaçant $\mathbf{p}' \rightarrow -\mathbf{p}'$,

$$[\phi_{\mathbf{p}}, \pi_{\mathbf{p}'}^\dagger] = i \mathcal{V} \delta_{\mathbf{p}, \mathbf{p}'} \quad \text{et} \quad [\phi_{\mathbf{p}}^\dagger, \pi_{\mathbf{p}'}] = i \mathcal{V} \delta_{\mathbf{p}, \mathbf{p}'} , \quad (2.54)$$

auquel on peut ajouter les relations triviales suivantes :

$$[\phi_{\mathbf{p}}, \phi_{\mathbf{p}'}] = [\pi_{\mathbf{p}}, \pi_{\mathbf{p}'}] = 0 \quad (2.55)$$

Lagrangien et transformées de Fourier Notons que le lagrangien (2.44) pourrait aussi d'exprimer en fonction des transformées de Fourier, avec le résultat suivant :

$$L = \frac{1}{2\mathcal{V}} \sum_{\mathbf{p}} [\dot{\phi}_{\mathbf{p}}^* \dot{\phi}_{\mathbf{p}} - \omega_{\mathbf{p}}^2 \phi_{\mathbf{p}}^* \phi_{\mathbf{p}}] \quad (2.56)$$

La raison pour laquelle nous sommes passés au hamiltonien directement dans l'espace réel au lieu de faire le passage après avoir exprimé le lagrangien comme ci-dessus est subtile : la transformée de Fourier $\phi_{\mathbf{p}}$ est une variable complexe, de plus contrainte par la relation $\phi_{\mathbf{p}}^* = \phi_{-\mathbf{p}}$. Le chemin qui mène de ce lagrangien vers un hamiltonien n'est donc pas tout à fait assuré. Il faudrait en fait restreindre la somme sur \mathbf{p} à la moitié des valeurs admissible, pour éviter de compter deux fois \mathbf{p} et $-\mathbf{p}$:

$$L = \frac{1}{\mathcal{V}} \sum_{\mathbf{p}^*} [\dot{\phi}_{\mathbf{p}}^* \dot{\phi}_{\mathbf{p}} - \omega_{\mathbf{p}}^2 \phi_{\mathbf{p}}^* \phi_{\mathbf{p}}] \quad (2.57)$$

où la notation \mathbf{p}^* signifie que la somme est prise sur un seul vecteur d'onde de la paire $\{\mathbf{p}, -\mathbf{p}\}$; le cas $\mathbf{p} = 0$ est manifestement spécial, et devrait conserver le facteur $\frac{1}{2}$, mais nous allons négliger ce détail dans ce qui suit. On doit alors traiter les variables $\phi_{\mathbf{p}}$ et $\phi_{\mathbf{p}}^*$ comme indépendantes, et les moments conjugués sont alors

$$\frac{\partial L}{\partial \dot{\phi}_{\mathbf{p}}} = \frac{1}{\mathcal{V}} \dot{\phi}_{\mathbf{p}}^* := \frac{1}{\mathcal{V}} \pi_{\mathbf{p}} \quad \text{et} \quad \frac{\partial L}{\partial \dot{\phi}_{\mathbf{p}}^*} = \frac{1}{\mathcal{V}} \dot{\phi}_{\mathbf{p}} := \frac{1}{\mathcal{V}} \pi_{\mathbf{p}}^* \quad (2.58)$$

ce qui mène aux crochets de Poisson suivants :

$$[\phi_{\mathbf{p}}, \pi_{\mathbf{p}'}]_{\text{P}} = \mathcal{V} \delta_{\mathbf{p}, \mathbf{p}'} \quad [\phi_{\mathbf{p}}^*, \pi_{\mathbf{p}'}^*]_{\text{P}} = \mathcal{V} \delta_{\mathbf{p}, \mathbf{p}'} \quad (2.59)$$

On constate que la définition de $\pi_{\mathbf{p}}$ coïncide avec celle trouvée plus haut. La définition habituelle du hamiltonien mène à

$$H = \frac{1}{\mathcal{V}} \sum_{\mathbf{p}^*} [\pi_{\mathbf{p}} \dot{\phi}_{\mathbf{p}} + \pi_{\mathbf{p}}^* \dot{\phi}_{\mathbf{p}}^*] - L \quad (2.60)$$

$$= \frac{1}{\mathcal{V}} \sum_{\mathbf{p}^*} [\frac{1}{2} \pi_{\mathbf{p}}^* \pi_{\mathbf{p}} + \omega_{\mathbf{p}}^2 \phi_{\mathbf{p}}^* \phi_{\mathbf{p}}] \quad (2.61)$$

$$= \frac{1}{2\mathcal{V}} \sum_{\mathbf{p}} [\pi_{\mathbf{p}}^* \pi_{\mathbf{p}} + \omega_{\mathbf{p}}^2 \phi_{\mathbf{p}}^* \phi_{\mathbf{p}}] \quad (2.62)$$

où on a utilisé le fait que $\pi_{\mathbf{p}}^* = \pi_{-\mathbf{p}}$. Donc en fin de compte on retrouve les mêmes résultats que ci-haut.

2.B.2 Opérateurs de création et d'annihilation

Nous pouvons alors introduire les opérateurs d'échelle :

$$a_{\mathbf{p}} = \sqrt{\frac{\omega_{\mathbf{p}}}{2\gamma}} \left(\phi_{\mathbf{p}} + i \frac{\pi_{\mathbf{p}}}{\omega_{\mathbf{p}}} \right) \quad a_{\mathbf{p}}^{\dagger} = \sqrt{\frac{\omega_{\mathbf{p}}}{2\gamma}} \left(\phi_{\mathbf{p}}^{\dagger} - i \frac{\pi_{\mathbf{p}}}{\omega_{\mathbf{p}}} \right) \quad (2.63)$$

Notons que, contrairement à $\phi_{\mathbf{p}}$ et $\pi_{\mathbf{p}}$, on a $a_{\mathbf{p}}^{\dagger} \neq a_{-\mathbf{p}}$. Ces opérateurs obéissent aux relations de commutation $[a_{\mathbf{p}}, a_{\mathbf{q}}^{\dagger}] = \delta_{\mathbf{p}\mathbf{q}}$. En effet,

$$[a_{\mathbf{p}}, a_{\mathbf{q}}^{\dagger}] = -i \frac{1}{2} \sqrt{\frac{\omega_{\mathbf{p}}}{\omega_{\mathbf{q}}\gamma}} [\phi_{\mathbf{p}}, \pi_{\mathbf{q}}^{\dagger}] + i \frac{1}{2} \sqrt{\frac{\omega_{\mathbf{q}}}{\omega_{\mathbf{p}}\gamma}} [\pi_{\mathbf{p}}, \phi_{\mathbf{q}}^{\dagger}] \quad (2.64)$$

$$= \frac{1}{2} \sqrt{\frac{\omega_{\mathbf{p}}}{\omega_{\mathbf{q}}}} \delta_{\mathbf{p},\mathbf{q}} + \frac{1}{2} \sqrt{\frac{\omega_{\mathbf{q}}}{\omega_{\mathbf{p}}}} \delta_{\mathbf{p},\mathbf{q}} \quad (2.65)$$

$$= \delta_{\mathbf{p},\mathbf{q}} \quad (2.66)$$

D'autre part,

$$[a_{\mathbf{p}}, a_{\mathbf{q}}] = i \frac{1}{2} \sqrt{\frac{\omega_{\mathbf{p}}}{\omega_{\mathbf{q}}\gamma}} [\phi_{\mathbf{p}}, \pi_{\mathbf{q}}] + i \frac{1}{2} \sqrt{\frac{\omega_{\mathbf{q}}}{\omega_{\mathbf{p}}\gamma}} [\pi_{\mathbf{p}}, \phi_{\mathbf{q}}] \quad (2.67)$$

$$= -\frac{1}{2} \sqrt{\frac{\omega_{\mathbf{p}}}{\omega_{\mathbf{q}}}} \delta_{\mathbf{p},-\mathbf{q}} + \frac{1}{2} \sqrt{\frac{\omega_{\mathbf{q}}}{\omega_{\mathbf{p}}}} \delta_{\mathbf{p},-\mathbf{q}} \quad (2.68)$$

$$= 0 \quad (2.69)$$

Calculons ce qui devrait être l'opérateur du nombre pour l'oscillateur harmonique :

$$N_{\mathbf{p}} = a_{\mathbf{p}}^{\dagger} a_{\mathbf{p}} = \frac{1}{\gamma} \left[\frac{\omega_{\mathbf{p}}}{2} \phi_{\mathbf{p}}^{\dagger} \phi_{\mathbf{p}} + \frac{1}{2\omega_{\mathbf{p}}} \pi_{\mathbf{p}}^{\dagger} \pi_{\mathbf{p}} + \frac{i}{2} \phi_{\mathbf{p}}^{\dagger} \pi_{\mathbf{p}} - \frac{i}{2} \pi_{\mathbf{p}}^{\dagger} \phi_{\mathbf{p}} \right] \quad (2.70)$$

$$= \frac{1}{\gamma} \left[\frac{1}{2} \omega_{\mathbf{p}} \phi_{\mathbf{p}}^{\dagger} \phi_{\mathbf{p}} + \frac{1}{2\omega_{\mathbf{p}}} \pi_{\mathbf{p}}^{\dagger} \pi_{\mathbf{p}} + \frac{i}{2} \phi_{\mathbf{p}}^{\dagger} \pi_{\mathbf{p}} - \frac{i}{2} \phi_{\mathbf{p}} \pi_{\mathbf{p}}^{\dagger} \right] - \frac{1}{2} \quad (2.71)$$

$$= \frac{1}{\gamma} \left[\frac{1}{2} \omega_{\mathbf{p}} \phi_{\mathbf{p}}^{\dagger} \phi_{\mathbf{p}} + \frac{1}{2\omega_{\mathbf{p}}} \pi_{\mathbf{p}}^{\dagger} \pi_{\mathbf{p}} + \frac{i}{2} \phi_{-\mathbf{p}} \pi_{\mathbf{p}} - \frac{i}{2} \phi_{\mathbf{p}} \pi_{-\mathbf{p}} \right] - \frac{1}{2} \quad (2.72)$$

En combinant \mathbf{p} et $-\mathbf{p}$ et en divisant par deux, on trouve

$$\frac{1}{2} a_{\mathbf{p}}^{\dagger} a_{\mathbf{p}} + \frac{1}{2} a_{-\mathbf{p}}^{\dagger} a_{-\mathbf{p}} = \frac{1}{\gamma} \left[\frac{1}{2} \omega_{\mathbf{p}} \phi_{\mathbf{p}}^{\dagger} \phi_{\mathbf{p}} + \frac{1}{2\omega_{\mathbf{p}}} \pi_{\mathbf{p}}^{\dagger} \pi_{\mathbf{p}} \right] - \frac{1}{2} \quad (2.73)$$

En multipliant par $\omega_{\mathbf{p}} = \omega_{-\mathbf{p}}$ et en sommant sur \mathbf{p} , on trouve enfin

$$\sum_{\mathbf{p}} \omega_{\mathbf{p}} (a_{\mathbf{p}}^{\dagger} a_{\mathbf{p}} + \frac{1}{2}) = \frac{1}{\gamma} \sum_{\mathbf{p}} \left[\frac{1}{2} \omega_{\mathbf{p}}^2 \phi_{\mathbf{p}}^{\dagger} \phi_{\mathbf{p}} + \frac{1}{2} \pi_{\mathbf{p}}^{\dagger} \pi_{\mathbf{p}} \right] \quad (2.74)$$

ce qui constitue effectivement le hamiltonien (2.50) du système :

$$H = \sum_{\mathbf{p}} \omega_{\mathbf{p}} (a_{\mathbf{p}}^{\dagger} a_{\mathbf{p}} + \frac{1}{2}) \quad (2.75)$$

Enfin, $\phi_{\mathbf{p}}$ peut s'exprimer ainsi en fonction des opérateurs d'échelle :

$$\phi_{\mathbf{p}} = \sqrt{\frac{\gamma}{2\omega_{\mathbf{p}}}} (a_{\mathbf{p}} + a_{-\mathbf{p}}^{\dagger}) \quad (2.76)$$

et donc

$$\phi(\mathbf{r}) = \frac{1}{\sqrt{\gamma}} \sum_{\mathbf{p}} \frac{1}{\sqrt{2\omega_{\mathbf{p}}}} (a_{\mathbf{p}} + a_{-\mathbf{p}}^{\dagger}) e^{i\mathbf{p}\cdot\mathbf{r}} \quad (2.77)$$

En faisant le changement $\mathbf{p} \rightarrow -\mathbf{p}$ dans le deuxième terme de la somme, on trouve finalement :

$$\phi(\mathbf{r}) = \frac{1}{\sqrt{\gamma}} \sum_{\mathbf{p}} \frac{1}{\sqrt{2\omega_{\mathbf{p}}}} (a_{\mathbf{p}} e^{i\mathbf{p}\cdot\mathbf{r}} + a_{\mathbf{p}}^{\dagger} e^{-i\mathbf{p}\cdot\mathbf{r}}) \quad (2.78)$$

2.B.3 Particules en tant que quanta du champ

Comme le système n'est qu'une collection d'oscillateurs harmoniques, on connaît bien les états propres de l'énergie. Les états propres de chaque oscillateur sont notés $|n_{\mathbf{q}}\rangle$ ($n_{\mathbf{q}}$ un entier positif ou nul), avec les propriétés :

$$a_{\mathbf{q}}|n_{\mathbf{q}}\rangle = \sqrt{n_{\mathbf{q}}}|n_{\mathbf{q}}-1\rangle \quad a_{\mathbf{q}}^{\dagger}|n_{\mathbf{q}}\rangle = \sqrt{n_{\mathbf{q}}+1}|n_{\mathbf{q}}+1\rangle \quad a_{\mathbf{q}}^{\dagger}a_{\mathbf{q}}|n_{\mathbf{q}}\rangle = n_{\mathbf{q}}|n_{\mathbf{q}}\rangle \quad (2.79)$$

L'état fondamental est $|0\rangle$ et l'énergie de l'état $|n_{\mathbf{q}}\rangle$ est $n_{\mathbf{q}}\hbar\omega_{\mathbf{q}}$, par rapport à celle de l'état fondamental. L'opérateur $N_{\mathbf{q}} = a_{\mathbf{q}}^{\dagger}a_{\mathbf{q}}$ prend des valeurs propres entières, est appelé l'*opérateur du nombre*. Le hamiltonien s'exprime ensuite comme $H = \sum_{\mathbf{q}} \omega_{\mathbf{q}} N_{\mathbf{q}}$ (modulo une constante additive).

L'état fondamental du système complet, qu'on notera $|0\rangle$, est le produit tensoriel des états fondamentaux de chaque oscillateur :

$$|0\rangle = |0\rangle_1 \otimes |0\rangle_2 \otimes |0\rangle_3 \otimes \cdots \otimes |0\rangle_N \quad (2.80)$$

où $|0\rangle_{\mathbf{q}}$ désigne le fondamental de l'oscillateur n° q .

Les états excités sont obtenus en appliquant les opérateurs d'échelle $a_{\mathbf{q}}^{\dagger}$ de toutes les manières possibles. En particulier, l'état excité le plus simple est obtenu en excitant l'un des oscillateurs seulement : $a_{\mathbf{q}}^{\dagger}|0\rangle$. Cette excitation se propage avec une quantité de mouvement \mathbf{q} et une énergie $\omega_{\mathbf{q}}$ et est interprétée comme une particule. L'opérateur $N_{\mathbf{q}} = a_{\mathbf{q}}^{\dagger}a_{\mathbf{q}}$ est alors le nombre de particules possédant une quantité de mouvement \mathbf{q} . Comme l'opérateur $a_{\mathbf{q}}^{\dagger}$ augmente de 1 la valeur propre de ce nombre, on dit qu'il crée une particule et on l'appelle opérateur de *création*. De même, l'opérateur $a_{\mathbf{q}}$ diminue de 1 la valeur de $n_{\mathbf{q}}$ et est appelé *opérateur d'annihilation* ou de *destruction* de particules.

Les particules ainsi créées sont des *bosons* : il est possible d'en mettre autant qu'on veut dans un mode de vecteur d'onde donné : l'état suivant

$$\frac{1}{\sqrt{n!}} (a_{\mathbf{q}}^{\dagger})^n |0\rangle \quad (2.81)$$

représente un état où n particules occupent le mode de vecteur d'onde \mathbf{q} . D'autre part, comme les opérateurs de création et d'annihilation associés à des modes différents commutent entre eux, l'état quantique obtenu par application d'opérateurs de création ne dépend pas de l'ordre dans lequel ces opérateurs sont appliqués (c.-à-d. le signe est le même dans tous les cas).

C Interactions et diagrammes de Feynman

2.C.1 Hamiltonien d'interaction

Modifions maintenant notre modèle simple de billes en ajoutant un terme anharmonique proportionnel à u_r^3 l'énergie potentielle qui figure dans le lagrangien (2.1) :

$$L = \frac{1}{2}\mu \sum_{r=1}^N \{ \dot{u}_r^2 - \Omega^2 u_r^2 - \Gamma^2 (u_r - u_{r+1})^2 - \lambda u_r^3 \} \quad (2.82)$$

Nous allons considérer cet ajout comme une perturbation H_1 sur le hamiltonien décrivant les particules indépendantes. En fonction du champ ϕ , cette perturbation s'écrit

$$H_1 = \frac{g}{6} \int dx \phi(x)^3 \quad \text{où} \quad g := 6\lambda \sqrt{\frac{a}{\mu}} \quad (2.83)$$

(la motivation pour le facteur $\frac{1}{6}$ apparaîtra sous peu).

Si on exprime le champ en fonction des opérateurs de création et d'annihilation comme à l'éq. (2.78), on trouve, en dimension 1,

$$\begin{aligned} H_1 &= \frac{g}{6} \frac{1}{(2L)^{3/2}} \sum_{p,p',q} \int dx \frac{1}{\sqrt{\omega_{p'}\omega_q\omega_p}} (a_p e^{ipx} + a_p^\dagger e^{-ipx}) (a_{p'} e^{ip'x} + a_{p'}^\dagger e^{-ip'x}) (a_q e^{iqx} + a_q^\dagger e^{-iqx}) \\ &= \frac{g}{6} \frac{1}{(2L)^{3/2}} \sum_{p,p',q} \int dx \frac{1}{\sqrt{\omega_{p'}\omega_q\omega_p}} [a_p a_{p'} a_q e^{ix(p+p'+q)} + a_p a_{p'} a_q^\dagger e^{ix(p+p'-q)} + a_p a_{p'}^\dagger a_q e^{ix(p-p'+q)} \\ &\quad + a_p a_{p'}^\dagger a_q^\dagger e^{ix(p-p'-q)} + a_p^\dagger a_{p'} a_q e^{ix(-p+p'+q)} + a_p^\dagger a_{p'} a_q^\dagger e^{ix(-p+p'-q)} \\ &\quad + a_p^\dagger a_{p'}^\dagger a_q e^{ix(-p-p'+q)} + a_p^\dagger a_{p'}^\dagger a_q^\dagger e^{-ix(p+p'+q)}] \\ &= \frac{g}{6} \frac{1}{2(2L)^{1/2}} \sum_{p,p',q} \frac{1}{\sqrt{\omega_{p'}\omega_q\omega_p}} [a_p a_{p'} a_q \delta_{p+p'+q} + a_p a_{p'} a_q^\dagger \delta_{p+p'-q} + a_p a_{p'}^\dagger a_q \delta_{p-p'+q} \\ &\quad + a_p a_{p'}^\dagger a_q^\dagger \delta_{p-p'-q} + a_p^\dagger a_{p'} a_q \delta_{-p+p'+q} + a_p^\dagger a_{p'} a_q^\dagger \delta_{-p+p'-q} \\ &\quad + a_p^\dagger a_{p'}^\dagger a_q \delta_{-p-p'+q} + a_p^\dagger a_{p'}^\dagger a_q^\dagger \delta_{p+p'+q}] \\ &= \frac{g}{6} \frac{1}{2(2L)^{1/2}} \sum_{p,q} \frac{1}{\sqrt{\omega_{p+q}\omega_q\omega_p}} [a_p a_{-p-q} a_q + a_{-p} a_{q+p} a_q^\dagger + a_p a_{q+p}^\dagger a_q + a_p^\dagger a_{p+q} a_{-q} \\ &\quad + a_p a_{p+q}^\dagger a_{-q}^\dagger + a_p^\dagger a_{p+q} a_q^\dagger + a_{-p}^\dagger a_{p+q}^\dagger a_q + a_p^\dagger a_{-p-q}^\dagger a_q^\dagger] \end{aligned} \quad (2.84)$$

Quelques explications sur ce qui précède : Nous devons utiliser trois variables de sommation distinctes (p, p', q) puisque l'interaction comporte trois facteurs. Pour passer de la deuxième à la troisième équation, nous avons intégré sur x , ce qui a généré des deltas de Kronecker sur les impulsions. Pour arriver à la dernière équation, nous avons tiré profit de ces deltas de Kronecker pour éliminer la somme sur p' , en posant $p' = p + q$ ou $p' = -p - q$, de manière à conserver le même préfacteur $\omega_{p+q}^{-1/2}$ pour tous les termes. Pour ce faire, nous avons changé la sommation de p à $-p$ ou de q à $-q$ au besoin. En procédant à quelques changements de variable d'intégration supplémentaires, on arrive à la forme équivalente suivante :

$$H_1 = \frac{g}{6} \frac{1}{2(2L)^{1/2}} \sum_{p,q} \frac{1}{\sqrt{\omega_{p+q} \omega_q \omega_p}} \left[a_p a_{-p-q} a_q + a_{q+p}^\dagger a_p a_q + a_p a_{q+p}^\dagger a_q + a_p a_q a_{q+p}^\dagger + \text{c.h.} \right] \quad (2.85)$$

Remarquons que chacun des termes de cette expression conserve la quantité de mouvement, c'est-à-dire ne la change pas. Par exemple, l'effet du premier terme ($a_p a_{-p-q} a_q$) est de détruire trois particules, d'impulsions $q, -p-q$ et p respectivement, ce qui ne change pas l'impulsion totale du système. Par contre, les trois termes qui suivent créent une particule d'impulsion $p+q$ et en détruisent deux d'impulsions p et q , ce qui ne change pas non plus l'impulsion totale. Cette conservation de l'impulsion est attribuable à l'invariance par translation $r \rightarrow r+1$ du hamiltonien original, ou encore $x \rightarrow x+s$ dans la limite continue, en raison des conditions aux limites périodiques.

Plutôt que de travailler avec le modèle unidimensionnel simple que nous avons utilisé comme motivation à la théorie quantique des champs, nous allons passer directement à trois dimensions d'espace, avec le hamiltonien d'interaction suivant :

$$H_1 = \frac{g}{6} \int d^3x \phi(x)^3 \quad (2.86)$$

figurant dans le lagrangien complet

$$\mathcal{L} = \frac{1}{2} \partial_\mu \phi \partial^\mu \phi - \frac{1}{2} m^2 \phi^2 - \frac{1}{6} g \phi^3 \quad (2.87)$$

La forme ci-dessus du terme d'interaction se généralise immédiatement au cas d'une théorie formulée en trois dimensions d'espace :

$$H_1 = \frac{g}{6} \frac{1}{2(2\mathcal{V})^{1/2}} \sum_{\mathbf{p},\mathbf{q}} \frac{1}{\sqrt{\omega_{\mathbf{p}+\mathbf{q}} \omega_{\mathbf{q}} \omega_{\mathbf{p}}}} \left[a_{\mathbf{p}} a_{-\mathbf{p}-\mathbf{q}} a_{\mathbf{q}} + a_{\mathbf{q}+\mathbf{p}}^\dagger a_{\mathbf{p}} a_{\mathbf{q}} + a_{\mathbf{p}} a_{\mathbf{q}+\mathbf{p}}^\dagger a_{\mathbf{q}} + a_{\mathbf{p}} a_{\mathbf{q}} a_{\mathbf{q}+\mathbf{p}}^\dagger + \text{c.h.} \right] \quad (2.88)$$

2.C.2 Diffusion entre particules : exemple

Le terme d'interaction (2.88) fait que les particules définies en tant qu'états propres du hamiltonien (2.75) ne sont plus des états propres du hamiltonien complet. En pratique, cela signifie, par exemple, qu'un état propre du hamiltonien non perturbé comportant deux particules développera, en fonction du temps, une composante le long d'autres états propres, certains comportant deux particules, d'autres même plus.

La théorie des perturbations nous permet de calculer une probabilité de transition par unité de temps d'un état initial $|i\rangle$ vers un état final $|f\rangle$, tous deux états propres du hamiltonien non perturbé (2.75). En particulier, nous pouvons ainsi calculer une section différentielle de diffusion. Rappelons la règle d'or de Fermi, qui donne le taux de transition d'un état initial $|i\rangle$ vers un état final $|f\rangle$:

$$\omega_{i \rightarrow f} = 2\pi |M|^2 \delta(E_f - E_i) \quad (2.89)$$

où l'amplitude M est, aux deux ordres les plus bas de la théorie des perturbations,

$$M = \langle f | H_1 | i \rangle + \sum_n \frac{\langle f | H_1 | n \rangle \langle n | H_1 | i \rangle}{E_i - E_n} + \dots \quad (2.90)$$

où le deuxième terme est une somme sur les états propres $|n\rangle$ du hamiltonien non perturbé, qualifiés d'*états virtuels*, dont les énergies sont E_n . Le deuxième terme doit être considéré si le premier s'annule.

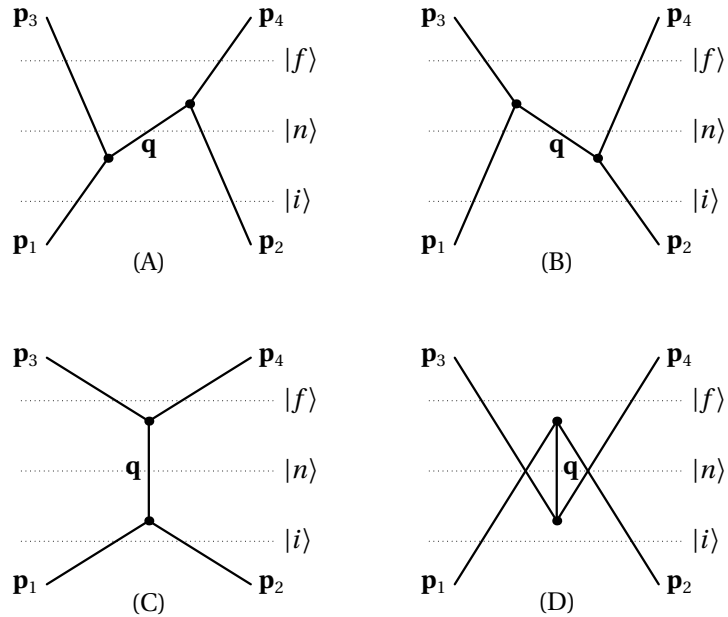


FIGURE 2.3

Considérons maintenant l'exemple suivant. Deux particules, d'impulsions \mathbf{p}_1 et \mathbf{p}_2 , entrent en collision en vertu de l'interaction (2.88), pour donner dans l'état final deux particules d'impulsions \mathbf{p}_3 et \mathbf{p}_4 . Comme l'énergie et la quantité de mouvement sont conservées lors de ce processus, on a

$$\omega_1 + \omega_2 = \omega_3 + \omega_4 \quad \text{et} \quad \mathbf{p}_1 + \mathbf{p}_2 = \mathbf{p}_3 + \mathbf{p}_4 \quad (2.91)$$

Les états initial et final sont

$$|i\rangle = a_1^\dagger a_2^\dagger |0\rangle \quad |f\rangle = a_3^\dagger a_4^\dagger |0\rangle \quad (2.92)$$

où nous employons la notation abrégée $a_3 = a_{\mathbf{p}_3}$, $a_{1-3} = a_{\mathbf{p}_1-\mathbf{p}_3}$, etc.

Le hamiltonien d'interaction (2.88) ne permet pas de passer directement de l'état initial à l'état final, c'est-à-dire que $\langle f | H_1 | i \rangle = 0$. Cela est dû au fait que H_1 doit nécessairement changer le nombre de

particules, par ± 1 ou ± 3 . Donc l'amplitude de diffusion M est nulle au premier ordre de la théorie des perturbations et on doit considérer les termes du deuxième ordre.

Les états intermédiaires possibles contiennent forcément une particule ou trois particules. L'interaction conserve cependant la quantité de mouvement, ce qui fait que les états intermédiaires doivent avoir la même quantité de mouvement que les états initial ou final. Nous ne sommes intéressés que par les états intermédiaires $|n\rangle$ tels que $\langle f|H_1|n\rangle\langle n|H_2|i\rangle$ est non nul. Il n'existe que six possibilités, modulo certaines permutations des termes d'interaction. Quatre de ces possibilités sont représentées sur la figure 2.3 par des diagrammes illustrant chaque état intermédiaire impliqué. Les deux autres sont obtenues en échangeant \mathbf{p}_3 et \mathbf{p}_4 dans les cas (A) et (B). Dans ces diagrammes, les particules initiales apparaissent en bas et les particules finales en haut, et les particules se «propagent» du bas vers le haut, en suivant les traits dessinés. Il ne s'agit pas strictement d'une propagation dans le temps, mais plutôt d'une succession d'apparitions des différents états, de la droite vers la gauche, dans le deuxième terme de (2.90).

Expliquons, en commençant par le premier diagramme (cas (A) sur la figure). Les trois états (initial, intermédiaire et final) sont représentés successivement de bas en haut. Un sommet (ou vertex) du diagramme correspond à l'action du hamiltonien d'interaction H_1 , qui nous fait passer successivement de l'état $|i\rangle$ à l'état $|n\rangle = a_3^\dagger a_2^\dagger a_{1-3}^\dagger |0\rangle$, à l'état $|f\rangle$. Le premier vertex représente l'action du terme $a_3^\dagger a_{1-3}^\dagger a_1$ de H_1 , ainsi que des termes obtenus des 5 autres permutations des mêmes facteurs, qui détruisent la particule d'impulsion \mathbf{p}_1 et la remplacent par une particule d'impulsion \mathbf{p}_3 , en créant une particule intermédiaire d'impulsion $\mathbf{q} = \mathbf{p}_1 - \mathbf{p}_3$. On associe à ce vertex l'élément de matrice

$$\langle n|H_1|i\rangle = g \frac{1}{\sqrt{8\mathcal{V}\omega_1\omega_3\omega_{1-3}}} \quad (2.93)$$

Les 6 permutations de ce terme ont compensé le facteur $\frac{1}{6}$ qui multiplie la constante de couplage g dans H_1 . Le deuxième vertex représente l'action de $a_4^\dagger a_{1-3} a_2$ dans H_1 (et de ses permutations), qui détruisent la particule d'impulsion \mathbf{p}_2 et la remplacent par une particule d'impulsion \mathbf{p}_4 , tout en annihilant la particule intermédiaire. On associe à ce vertex l'élément de matrice

$$\langle f|H_1|n\rangle = g \frac{1}{\sqrt{8\mathcal{V}\omega_2\omega_4\omega_{1-3}}} \quad (2.94)$$

le dénominateur figurant dans l'amplitude (2.90) est

$$E_i - E_n = (\omega_1 + \omega_2) - (\omega_2 + \omega_3 + \omega_{1-3}) = \omega_1 - \omega_3 - \omega_{1-3} \quad (2.95)$$

Au total, ce diagramme a la valeur numérique suivante :

$$M^{(A)} = \frac{g^2}{8\mathcal{V}\sqrt{\omega_1\omega_2\omega_3\omega_4}} \frac{1}{\omega_{1-3}} \frac{1}{\omega_1 - \omega_3 - \omega_{1-3}} \quad (2.96)$$

Pour calculer la contribution du diagramme (B), il suffit d'échanger $1 \leftrightarrow 2$ et $3 \leftrightarrow 4$, tout en gardant à l'esprit que $\omega_1 - \omega_3 = \omega_4 - \omega_2$, $\mathbf{p}_1 - \mathbf{p}_3 = \mathbf{p}_4 - \mathbf{p}_2$ et que $\omega_{3-1} = \omega_{1-3}$. On obtient alors

$$M^{(B)} = \frac{g^2}{8\mathcal{V}\sqrt{\omega_1\omega_2\omega_3\omega_4}} \frac{1}{\omega_{1-3}} \frac{1}{\omega_3 - \omega_1 - \omega_{1-3}} \quad (2.97)$$

En combinant ces deux amplitudes, on obtient

$$M^{(A+B)} = \frac{g^2}{4\mathcal{V}\sqrt{\omega_1\omega_2\omega_3\omega_4}} \frac{1}{(\omega_1 - \omega_3)^2 - \omega_{1-3}^2} = \frac{g^2}{4\mathcal{V}\sqrt{\omega_1\omega_2\omega_3\omega_4}} \frac{1}{(p_1 - p_3)^2 - m^2} \quad (2.98)$$

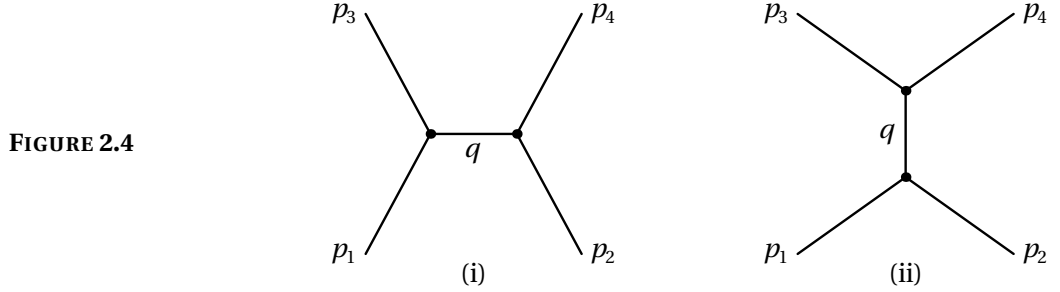
où p_1 et p_3 désignent les quadrivecteurs énergie-impulsion des particules correspondantes. La dernière égalité provient du fait que $\omega_{1-3}^2 = (\mathbf{p}_1 - \mathbf{p}_3)^2 + m^2$.

On montre sans peine que les diagrammes (C) et (D) valent respectivement

$$\begin{aligned} M^{(C)} &= \frac{g^2}{8\mathcal{V}\sqrt{\omega_1\omega_2\omega_3\omega_4}} \frac{1}{\omega_{1+2}} \frac{1}{\omega_1 + \omega_2 - \omega_{1+2}} \\ M^{(D)} &= \frac{g^2}{8\mathcal{V}\sqrt{\omega_1\omega_2\omega_3\omega_4}} \frac{1}{\omega_{1+2}} \frac{1}{-\omega_1 - \omega_2 - \omega_{1+2}} \end{aligned} \quad (2.99)$$

et que leur somme vaut

$$M^{(C+D)} = \frac{g^2}{4\mathcal{V}\sqrt{\omega_1\omega_2\omega_3\omega_4}} \frac{1}{(p_1 + p_2)^2 - m^2} \quad (2.100)$$



Notons tout de suite que ces amplitudes ont été calculées en utilisant une normalisation discrète des états. Dans la normalisation relativiste, on devrait premièrement supprimer le facteur de volume, et ensuite multiplier par le facteur de normalisation $\sqrt{2\omega_i}$ pour chaque particule i figurant dans l'état initial ou final. On obtiendrait alors les combinaisons suivantes, qui sont invariantes de Lorentz :

$$\begin{aligned} \mathcal{M}^{(A+B)} &= \mathcal{M}^{(i)} = \frac{g^2}{(p_1 - p_3)^2 - m^2} \\ \mathcal{M}^{(C+D)} &= \mathcal{M}^{(ii)} = \frac{g^2}{(p_1 + p_2)^2 - m^2} \end{aligned} \quad (2.101)$$

Ces combinaisons sont représentées graphiquement par les deux diagrammes de la figure 2.4, dits *diagrammes de Feynman*. Ceux-ci interviennent dans une forme légèrement différente de la théorie des perturbations (la *théorie covariante des perturbations*) qui ne distingue pas entre les deux états intermédiaires A et B (ou C et D). Elle fait plutôt intervenir des *particules virtuelles* qui sont émises et absorbées par les particules réelles présentes dans les états final et initial. Les particules virtuelles sont représentées par les lignes internes dans les diagrammes de la figure 2.4 et possèdent les quadri-impulsions $q = p_1 - p_3$ et $q = p_1 + p_2$ dans les diagrammes (i) et (ii) respectivement. La quadri-impulsion d'une particule virtuelle est déterminée par la conservation de la quadri-impulsion à chacun des vertex du diagramme, et cette conservation garantit la conservation globale

de la quadri-impulsion entre l'état initial et l'état final. Cependant, la particule virtuelle n'est pas sur la *couche de masse*, c'est-à-dire que $q^2 \neq m^2$.

La notion de *couche de masse* désigne l'hypersurface, dans l'espace des énergies-impulsions, définie par l'équation $p^2 = p_0^2 - \mathbf{p}^2 = m^2$. Cette hypersurface a l'apparence d'un hyperboloïde. Un quadrivecteur qui ne tombe pas sur cette hypersurface ne peut pas représenter la quadri-impulsion d'une particule observable. Dans le cas qui nous occupe, p_1 et p_3 appartiennent à des particules réelles et ces quadrivecteurs sont sur la couche de masse : $p_1^2 = m^2$ et $p_3^2 = m^2$; par contre, la différence $p_1 - p_3$ ne peut pas l'être : en se plaçant dans le référentiel du centre de masse du processus ($\mathbf{p}_1 + \mathbf{p}_2 = 0$), on sait que les énergies E_1 et E_3 sont les mêmes et donc que $p_1 - p_3 = (0, \mathbf{p}_1 - \mathbf{p}_3)$ et donc que $(p_1 - p_3)^2 = -(\mathbf{p}_1 - \mathbf{p}_3)^2 < 0$. Cela signifie que le quadrivecteur $p_1 - p_3$ est de genre «espace», alors que le quadrivecteur d'une particule massive ($m > 0$) est de genre «temps». Comme $(p_1 - p_3)^2$ est un invariant, l'argument ne dépend pas du référentiel utilisé pour faire le calcul. De même, la combinaison $p_1 + p_2$ n'est pas sur la couche de masse non plus : dans le référentiel du centre de masse, cette quantité vaut $(2E, 0)$ et donc $(p_1 + p_2)^2 = 4E^2 = 4(\mathbf{p}^2 + m^2)$, où \mathbf{p} est la quantité de mouvement de l'une des deux particules initiales dans ce référentiel. Ceci est toujours plus grand que m^2 .

Résumons l'interprétation physique des calculs que nous venons de faire. Dans le cas parfaitement harmonique ($g = 0$), les états propres du système défini par le lagrangien (2.44) sont des particules de masse m , respectant la relation de dispersion $p_\mu p^\mu = m^2$. Si $g \neq 0$, le lagrangien contient en outre un terme cubique ($g\phi^3$) qui rend les équations de Lagrange correspondantes non linéaires. Ce terme cause une interaction entre les particules, qui ne sont plus désormais des états propres exacts. Si g n'est pas trop grand, on peut calculer l'effet de cette interaction en théorie des perturbations. Cette dernière peut être formulée en fonction de l'échange de «particules virtuelles» entre particules réelles. Les particules virtuelles sont bien sûr indétectables et ne sont pas sur la couche de masse. Cette interprétation s'applique à toutes les interactions fondamentales, mais est tributaire de la théorie des perturbations ; elle perd son sens si l'interaction est trop forte pour que la théorie des perturbations soit applicable.

2.C.3 Calcul de la section efficace

Nous allons indiquer comment compléter le calcul de la section différentielle de diffusion en se basant sur l'amplitude calculée ci-dessus. L'amplitude complète comprend les parties $\mathcal{M}^{(i)}$ et $\mathcal{M}^{(ii)}$, plus les termes obtenus en échangeant les rôles des deux particules finales ($3 \leftrightarrow 4$) dans $\mathcal{M}^{(i)}$. On obtient alors

$$\mathcal{M} = g^2 \left\{ \frac{1}{(p_1 - p_3)^2 - m^2} + \frac{1}{(p_1 - p_4)^2 - m^2} + \frac{1}{(p_1 + p_2)^2 - m^2} \right\} \quad (2.102)$$

Calculons la section différentielle de diffusion dans le référentiel du centre de masse. Posons donc

$$\mathbf{p}_1 = p\mathbf{x} \quad \mathbf{p}_2 = -p\mathbf{x} \quad \mathbf{p}_3 = p\mathbf{n} \quad \mathbf{p}_4 = -p\mathbf{n} \quad (2.103)$$

où \mathbf{n} est le vecteur unitaire dans la direction de la particule diffusée, faisant un angle θ avec l'axe des x . Les quatre énergies ω_i sont égales à l'énergie E du faisceau.

Dans le premier terme de l'amplitude,

$$\mathbf{q} = \mathbf{p}_1 - \mathbf{p}_3 = p(\mathbf{x} - \mathbf{n}) \implies (p_1 - p_3)^2 = -\mathbf{q}^2 = -p^2(2 - 2\cos\theta) = -4p^2 \sin^2 \theta / 2 \quad (2.104)$$

Dans le second,

$$\mathbf{q} = \mathbf{p}_1 - \mathbf{p}_4 = p(\mathbf{x} + \mathbf{n}) \implies (p_1 - p_4)^2 = -\mathbf{q}^2 = -p^2(2 + 2\cos\theta) = -4p^2 \cos^2 \theta / 2 \quad (2.105)$$

Dans le troisième,

$$\mathbf{q} = \mathbf{p}_1 + \mathbf{p}_2 = 0 \implies (p_1 + p_2)^2 = (\omega_1 + \omega_2)^2 = 4E^2 \quad (2.106)$$

L'amplitude totale est donc, dans ce référentiel,

$$\mathcal{M} = -g^2 \left\{ \frac{1}{4p^2 \sin^2 \theta / 2 + m^2} + \frac{1}{4p^2 \cos^2 \theta / 2 + m^2} - \frac{1}{4E^2 - m^2} \right\} \quad (2.107)$$

On calcule (voir problème 1.6) que dans le référentiel du centre de masse,

$$\frac{d\sigma}{d\Omega} = \frac{1}{2} \frac{|\mathcal{M}|^2}{(8\pi)^2} \frac{|\mathbf{p}_3|}{|\mathbf{p}_1|} \frac{1}{E^2} \quad (2.108)$$

où E est l'énergie totale des particules incidentes. Le préfacteur $\frac{1}{2}$ devant cette expression tient compte de l'identité des particules : il faut éviter, dans le décompte des états finaux, de compter comme deux états distincts les configurations qui s'obtiennent l'une de l'autre par échange de particules identiques. Dans le cas qui nous occupe, la section différentielle serait, si on l'exprime en fonction du facteur $\gamma = E/m$ du faisceau,

$$\frac{d\sigma}{d\Omega} = \frac{g^4}{128\pi^2 m^6 \gamma^2} \left(\frac{4\gamma^2 - 2}{2\gamma^4 - 1 - 2(\gamma^2 - 1)^2 \cos(4\theta)} + \frac{1}{4\gamma^2 - 1} \right)^2 \quad (2.109)$$

2.C.4 Règles de Feynman

Nous avons effectué un calcul simple dans le cadre de la théorie des perturbations dépendant du temps. Il existe cependant une version améliorée de la théorie des perturbations, dont Feynman fut le pionnier et Freeman Dyson l'architecte. Cette «théorie covariante des perturbation» ne sera pas étudiée en détail dans ce cours. Par contre, elle peut être appliquée sans qu'il soit nécessaire d'en comprendre tous les replis, car elle est habituellement formulée en fonction de diagrammes de Feynman. Ses règles peuvent être résumées ainsi, si on l'applique à la théorie du champ scalaire :

1. Les règles suivantes visent à construire une expression mathématique pour $i\mathcal{M}$, \mathcal{M} étant l'amplitude invariante du processus dans la normalisation relativiste.
2. Pour un processus donné, identifier les états initial et final et dessiner les lignes correspondantes.
3. Construire les diagrammes possibles à l'ordre N de la théorie des perturbations en introduisant N vertex, auxquels sont reliés les lignes externes et des lignes internes (correspondant à des particules virtuelles) introduites au besoin.
4. Chaque ligne porte une certaine quadri-impulsion : notons-la p_i pour les lignes externes, et q_i pour les lignes internes. La direction de cette quadri-impulsion est arbitraire, c'est-à-dire affaire de convention. On peut faire une analogie avec les lois de Kirchhoff dans la théorie des circuits. La quadri-impulsion est conservée à chaque vertex.

5. À chaque vertex, insérer un facteur $-ig$. Ajouter un facteur $(2\pi)^4 \delta(k_1 + k_2 + k_3)$, où $k_{1,2,3}$ sont les quadri-impulsions *entrantes* (changer le signe s'il s'agit de quadri-impulsions sortantes).
6. À chaque ligne interne est associé un *propagateur*. Pour un boson virtuel, ce propagateur est

$$-i \frac{1}{q^2 - m^2}$$

où q est la 4-impulsion du boson virtuel.

7. Intégrer sur toute quadri-impulsion interne q_i :

$$\int \frac{d^4 q_i}{(2\pi)^4} \dots$$

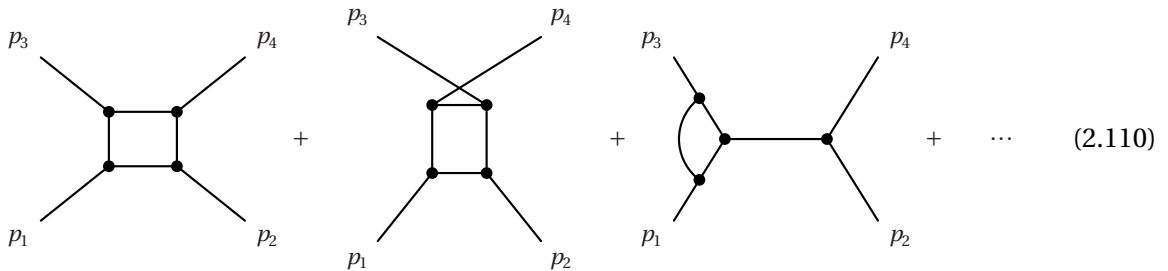
En pratique, si le diagramme ne contient pas de boucle fermée, les fonctions deltas qui imposent la conservation de la 4-impulsion à chaque vertex rendent ces intégrales triviales. Chaque boucle du diagramme correspond à une réelle intégration. Ces intégrales constituent la principale difficulté calculatoire, notamment en raison de leur caractère singulier.

8. La conservation de la quadri-impulsion entre les états initial et final est une conséquence de sa conservation à chaque vertex. Cependant, le facteur global $(2\pi)^4 \delta(p_1 + p_2 + \dots - p_n)$ associé à la conservation de l'énergie-impulsion doit être amputé, pour obtenir l'amplitude \mathcal{M} (par opposition à M). Voir à cet effet la discussion de la section 1.E.2).

Si on applique les règles ci-dessus au processus de diffusion étudié à la section 2.C.2 à l'ordre deux, on trouve trois diagrammes suivant : les deux diagrammes de la figure 2.4, ainsi qu'un troisième, obtenu du premier en permutant les indices 3 et 4. L'expression (2.102) pour l'amplitude totale en découle.

À l'ordre 4, le nombre de diagrammes est plus grand. Par exemple, en voici trois :

Prob. 2.4



$$(2.110)$$

2.C.5 Potentiel de Yukawa

Nous allons montrer dans cette section que l'interaction portée par les particules virtuelles correspond à un potentiel de la forme suivante :

$$U(r) = \frac{e^{-mr}}{r} . \quad (2.111)$$

Ce potentiel, dit *potentiel de Yukawa*, est semblable au potentiel coulombien en ce qu'il décroît comme l'inverse de la distance, mais diffère de ce dernier par un facteur exponentiel qui accentue la décroissance du potentiel, sur une distance caractéristique m^{-1} . Cette distance est appelée la *portée* du potentiel. La portée du potentiel est en fait la longueur d'onde réduite de Compton \hbar/mc de la particule médiatrice, si on restaure les unités habituelles.

Première méthode

Ajoutons au lagrangien (2.44) le terme suivant :

$$\mathcal{L} \rightarrow \mathcal{L} + \rho(\mathbf{r})\phi(\mathbf{r}) . \quad (2.112)$$

Si on retourne au lagrangien d'origine (2.1), ce terme revient à exercer une force constante $\rho(x_r)$ sur la bille située à la position x_r , un peu comme si une charge mécanique était attachée à la bille. Par le principe d'action-réaction, la bille, et donc le champ ϕ , va exercer une force égale et opposée sur la charge $\rho(x)$. Le hamiltonien du système serait modifié de la façon suivante :

$$H \rightarrow \int d^3r \left[\frac{1}{2}\pi(x)^2 + \frac{1}{2}m^2\phi^2 + \frac{1}{2}(\nabla\phi)^2 - \rho\phi \right] . \quad (2.113)$$

La stratégie est ici de déterminer comment l'ajout de cette charge modifie l'énergie du système. L'imposition d'une force constante sur chaque bille du système (2.1) va résulter en un déplacement de l'équilibre. Une nouvelle configuration d'équilibre va s'établir ; on peut la déterminer simplement en solutionnant les équations du mouvement en supposant une configuration stationnaire, c'est-à-dire indépendante du temps. Remarquons à cet effet que l'équation du mouvement du champ ϕ , suite à l'ajout de la charge ρ , est simplement l'équation de Klein-Gordon inhomogène :

$$(\partial_\mu\partial^\mu + m^2)\phi = -\rho \quad \text{ou, dans le cas stationnaire,} \quad -\nabla^2\phi + m^2\phi = -\rho . \quad (2.114)$$

En fonction des transformées de Fourier $\tilde{\phi}(\mathbf{p})$ et $\tilde{\rho}(\mathbf{p})$, cette dernière équation s'écrit simplement

$$(\mathbf{p}^2 + m^2)\tilde{\phi} = -\tilde{\rho} \implies \tilde{\phi} = -\frac{1}{\mathbf{p}^2 + m^2}\tilde{\rho} . \quad (2.115)$$

Quant à l'énergie de la configuration, elle est simplement donnée par

$$E[\rho] = \int d^3r \left[\frac{1}{2}m^2\phi^2 + \frac{1}{2}(\nabla\phi)^2 - \rho\phi \right] = \int \frac{d^3p}{(2\pi)^3} \left[\frac{1}{2}(\mathbf{p}^2 + m^2)\tilde{\phi}^*(\mathbf{p})\tilde{\phi}(\mathbf{p}) - \tilde{\rho}^*(\mathbf{p})\tilde{\phi}(\mathbf{p}) \right] , \quad (2.116)$$

où nous avons utilisé le théorème de Parseval pour passer à la dernière expression. En substituant ce que vaut $\tilde{\phi}$ en fonction de $\tilde{\rho}$, on trouve

$$E[\rho] = -\frac{1}{2} \int \frac{d^3p}{(2\pi)^3} \frac{\tilde{\rho}^*(\mathbf{p})\tilde{\rho}(\mathbf{p})}{\mathbf{p}^2 + m^2} = -\frac{1}{2} \int d^3r \, d^3r' \, \rho(\mathbf{r})\rho(\mathbf{r}') \int \frac{d^3p}{(2\pi)^3} \frac{1}{\mathbf{p}^2 + m^2} e^{i\mathbf{p}\cdot(\mathbf{r}-\mathbf{r}')} . \quad (2.117)$$

En définissant le potentiel

$$V(\mathbf{r}) = - \int \frac{d^3p}{(2\pi)^3} \frac{1}{\mathbf{p}^2 + m^2} e^{i\mathbf{p}\cdot\mathbf{r}} = -\frac{e^{-mr}}{r} = U(r) \quad r = |\mathbf{r}| \quad (2.118)$$

(voir la section 1.F.2 pour le calcul de la transformée de Fourier) on trouve l'expression finale

$$E[\rho] = \frac{1}{2} \int d^3r d^3r' \rho(\mathbf{r}) \rho(\mathbf{r}') U(|\mathbf{r} - \mathbf{r}'|) . \quad (2.119)$$

Cette dernière expression démontre clairement que la fonction $U(r)$ représente l'énergie potentielle d'interaction entre des charges séparées par une distance r , de sorte que l'énergie potentielle totale d'une distribution de charge est donnée par la somme des contributions provenant de toutes les paires de charges possibles (d'où le préfacteur $\frac{1}{2}$). En électrostatique, une expression semblable représente l'énergie potentielle électrique associée à une distribution de charge électrique. Le signe négatif dans $U(r)$ signifie bien sûr que l'interaction est attractive.

Le calcul que nous venons de présenter suppose que la distribution de charge ρ est statique. La force qui agit entre les phonons en vertu de l'interaction (2.88) est de même nature et décroît de même manière avec la distance. Cependant, elle est proportionnelle à la constante de couplage g et, plus important, elle est *retardée*, c'est-à-dire qu'elle prend un certain temps à se propager. Elle ne peut donc pas être décrite par un potentiel d'interaction qui ne dépend que de la distance. Ce serait en fait possible dans la limite où les particules en interaction sont lentes en comparaison des particules qui transmettent l'interaction, mais dans le cas qui nous occupe les particules qui subissent l'interaction et celles qui la transmettent sont de même nature. On remarque cependant l'intervention du facteur $(q^2 - m^2)^{-1}$ en relation avec la particule virtuelle dans les amplitudes (2.101). Ce facteur, qu'on appelle *propagateur*, devient, dans le cas statique, $-(\mathbf{q}^2 + m^2)$, ce qui est précisément la transformée de Fourier du potentiel de Yukawa.

Deuxième méthode

Une autre façon de comprendre l'émergence du potentiel de Yukawa est de considérer l'équation de Klein-Gordon en présence d'une source ponctuelle située à l'origine, dans la limite statique :

$$-\nabla^2 \phi + m^2 \phi = -g \delta(\mathbf{r}) . \quad (2.120)$$

Cette équation est l'analogue, pour un champ scalaire massif, de l'équation de Poisson pour le potentiel électrique Φ :

$$\nabla^2 \Phi = -e \delta(\mathbf{r}) \quad (2.121)$$

dont la solution est bien connue :

$$\Phi(\mathbf{r}) = \frac{e}{4\pi r} . \quad (2.122)$$

Dans le cas du champ scalaire, nous allons montrer que la solution est

$$\phi(\mathbf{r}) = -g \frac{e^{-mr}}{4\pi r} . \quad (2.123)$$

Pour ce faire, passons en coordonnées sphériques et remarquons que la solution ne doit pas dépendre des angles. L'équation se simplifie alors comme suit :

$$-\frac{1}{r} \frac{\partial^2}{\partial r^2} (r\phi) + m^2 \phi = -g \delta(\mathbf{r}) . \quad (2.124)$$

En dehors de l'origine et en fonction de la variable $u = r\phi$, cette équation différentielle devient

$$-\frac{d^2 u}{dr^2} + m^2 u = 0 \quad \text{dont la solution générale est} \quad u(r) = A e^{-mr} + B e^{mr} \quad (2.125)$$

où A et B sont des constantes. Afin d'en tirer une solution physique à énergie finie, on pose $B = 0$ et il reste à déterminer la constante A . Ceci se fait en intégrant l'équation différentielle (2.120) dans une sphère infinitésimale S de rayon ϵ centrée à l'origine. On trouve alors

$$-\int_S d^3 r \nabla \cdot \nabla \left(A \frac{e^{-mr}}{r} \right) + m^2 \int_S d^3 r A \frac{e^{-mr}}{r} = -g . \quad (2.126)$$

Le premier terme peut être intégré à l'aide du théorème de la divergence, ce qui donne

$$-\int_{\partial S} d\mathbf{a} \cdot \nabla \left(A \frac{e^{-mr}}{r} \right) + m^2 \int_S d^3 r A \frac{e^{-mr}}{r} = -g . \quad (2.127)$$

En développant le gradient et en supposant par la suite que $\epsilon \rightarrow 0$, on trouve

$$4\pi\epsilon^2 \left(A \frac{1}{r^2} + m A \frac{1}{r} \right) \Big|_{r=\epsilon} + 4\pi m^2 \int_0^\epsilon r^2 dr \frac{A}{r} = -g , \quad (2.128)$$

ce qui donne

$$4\pi A + 4\pi A m \epsilon + 2\pi A m^2 \epsilon^2 \rightarrow 4\pi A = -g \quad \text{ou encore} \quad A = -\frac{g}{4\pi} , \quad (2.129)$$

ce qui confirme la solution (2.123).

D Problèmes

Problème 2.1 Équation de Proca

Considérons une théorie dans laquelle figure un champ vectoriel A_μ , avec la densité lagrangienne suivante :

$$\mathcal{L} = -\frac{1}{4}F_{\mu\nu}F^{\mu\nu} + \frac{1}{2}m^2 A_\mu A^\mu \quad \text{où} \quad F_{\mu\nu} := \partial_\mu A_\nu - \partial_\nu A_\mu \quad (2.130)$$

A Montrez que le principe de moindre action appliqué à ce lagrangien entraîne l'équation différentielle suivante, dite *équation de Proca* :

$$\partial_\mu(\partial^\mu A^\nu - \partial^\nu A^\mu) + m^2 A^\nu = 0 \quad (2.131)$$

B En supposant une solution de type onde plane : $A^\nu = \epsilon^\nu e^{-ik_\mu x^\mu}$, où le quadrivecteur ϵ^ν est constant, montrez que le paramètre m a l'interprétation de la masse. Vous pouvez vous aider de Mathematica ou l'équivalent.

Problème 2.2 Fluctuations d'un champ scalaire

Considérons un champ quantique scalaire ϕ , dont le développement en modes est le suivant :

$$\phi(\mathbf{r}) = \sum_{\mathbf{q}} \frac{1}{\sqrt{2V\omega_q}} (a_{\mathbf{q}} e^{i\mathbf{q}\cdot\mathbf{r}} + a_{\mathbf{q}}^\dagger e^{-i\mathbf{q}\cdot\mathbf{r}}) \quad (2.132)$$

Imaginons que nous avons un moyen de « mesurer » ce champ en un endroit donné, mais que nos appareils sont entachés d'une certaine erreur, de sorte qu'ils ne mesurent pas le champ $\phi(\mathbf{r})$ directement, mais une certaine moyenne de ce champ autour d'une position donnée, qu'on définit comme

$$\Phi(\mathbf{r}) = \int d^3r' \phi(\mathbf{r}') f(\mathbf{r} - \mathbf{r}') \quad (2.133)$$

où f est une fonction qui ne dépend que de la distance, qu'on prend comme étant une gaussienne de largeur R :

$$f(r) = A \exp\left(-\frac{r^2}{R^2}\right) \quad A = \frac{1}{(\pi R^2)^{3/2}} \quad (2.134)$$

La constante A est choisie de manière à normaliser la fonction f à l'unité. Cette fonction caractérise l'imprécision de l'appareil de mesure et R est en quelque sorte la résolution de l'appareil.

A Montrez que la valeur moyenne dans le vide du « champ flou » $\Phi(0)$ est nulle.

B Calculez la fluctuation de ce champ dans le vide, c'est-à-dire sa variance, définie dans ce cas-ci comme $\Delta\Phi^2 = \langle 0|\Phi(0)^2|0\rangle$. Montrez que

$$\Delta\Phi^2 = \frac{1}{4\pi^2} \int_0^\infty dq \frac{q^2}{\omega_q} e^{-q^2 R^2/2} \quad (2.135)$$

C Montrez que cette variance tend vers l'infini quand $R \rightarrow 0$. Calculez l'intégrale dans les deux cas limites où (i) $m \gg 1/R$ et (ii) $m \ll 1/R$, m étant la masse du champ ϕ (le résultat est fini dans les deux cas).

Indice : n'hésitez pas à vous servir du théorème de Parseval et de la notion de convolution. Toutes les intégrales en jeu se font analytiquement, mais l'usage de tables ou de Mathematica est bien sûr permis.

Ceci démontre que les champs quantiques sont des quantités aux très fortes fluctuations...

Problème 2.3 Oscillations entre deux types de bosons

Une théorie impliquant deux champs scalaires ϕ_1 et ϕ_2 est définie par la densité lagrangienne suivante :

$$\mathcal{L} = \frac{1}{2} (\partial_\mu \phi_1 \partial^\mu \phi_1 + \partial_\mu \phi_2 \partial^\mu \phi_2 - m^2 \phi_1^2 - m^2 \phi_2^2) - \gamma \phi_1 \phi_2$$

A Expliquez comment exprimer cette théorie plus simplement, et décrivez son contenu physique. Indice : il est utile de concevoir les deux champs ϕ_1 et ϕ_2 comme les composantes d'un vecteur abstrait à deux composantes $\Phi = (\phi_1, \phi_2)$, et de transformer ce vecteur vers une autre base dans lequel il s'exprime à l'aide de deux autres champs scalaires χ_1 et χ_2 qui sont alors découplés.

B Trouver comment exprimer les opérateurs d'annihilation $a_{1\mathbf{k}}$ et $a_{2\mathbf{k}}$ des champs $\phi_{1,2}$ en fonction de ceux des champs $\chi_{1,2}$, qui seront notés $b_{1\mathbf{k}}$ et $b_{2\mathbf{k}}$. Exprimez ces nouveaux opérateurs dans la limite $\gamma \ll m$. Indice : exprimer le hamiltonien complet en fonction des opérateurs $a_{i\mathbf{k}}$ et $a_{i\mathbf{k}}^\dagger$. Des termes quadratiques en a et quadratiques en a^\dagger feront partie de l'expression. Afin de ramener le hamiltonien à une forme habituelle, il faut procéder à une transformation, dite de Bogolioubov, qui mélange les opérateurs de création et d'annihilation.

C Supposons maintenant que seul le champ ϕ_1 interagit avec les autres champs (ou lui-même), alors que ϕ_2 n'apparaît dans aucune interaction. Cela signifie que seules les particules associées au champ ϕ_1 peuvent être créées lors de collisions ou détectées. Supposons qu'un quantum du champ ϕ_1 créé à $x = 0$ se propage avec une quantité de mouvement p vers la droite, tout en étant non relativiste. Si un détecteur efficace à 100% est placé à une distance x de la source, obtenez une expression pour la probabilité $P(x)$ de détecter cette particule en fonction de x . Placez-vous toujours dans la limite $\gamma \ll m$. Indice : les quanta du champ ϕ_1 ne sont pas des états propres du hamiltonien, même en oubliant les interactions.

Problème 2.4 Diagrammes de Feynman à l'ordre 4 pour le champ scalaire

Dessinez tous les diagrammes contribuant à l'amplitude de diffusion boson-boson au quatrième ordre de la théorie des perturbations. Trois sont déjà illustrés à l'éq. (2.110).

Problème 2.5 Théorie en ϕ^4

Supposons que le terme d'interaction entre des particules décrites par un champ scalaire ϕ ait la forme $\frac{g}{4!}\phi^4$, au lieu de $\frac{g}{3!}\phi^3$. Le seul changement à apporter aux règles de Feynman est que quatre lignes doivent émerger d'un vertex au lieu de trois.

A Écrivez l'amplitude de diffusion \mathcal{M} pour le processus de collision comportant deux particules dans l'état initial et deux particules dans l'état final, à l'ordre le plus bas de la théorie des perturbations. Calculez ensuite la section différentielle de diffusion correspondante dans le référentiel du centre de masse (vous pouvez utiliser à cette fin la formule donnée dans le problème 1.6 sans la démontrer).

B Dessinez tous les diagrammes de Feynman qui corrigent cette amplitude à l'ordre suivant en théorie des perturbations.

Problème 2.6 Interaction entre deux types de bosons

Deux espèces de particules, de masses respectives m_1 et m_2 , sont décrites par des champs scalaires notés ϕ_1 et ϕ_2 . La densité lagrangienne décrivant leur comportement est définie par l'expression suivante :

$$\mathcal{L} = \frac{1}{2}(\partial_\mu \phi_1 \partial^\mu \phi_1 - m_1^2 \phi_1^2) + \frac{1}{2}(\partial_\mu \phi_2 \partial^\mu \phi_2 - m_2^2 \phi_2^2) - \frac{1}{2}g(\phi_1 \phi_2^2 + \phi_1^2 \phi_2) \quad (2.136)$$

Les règles de Feynman associées à cette théorie sont les mêmes que pour un champ scalaire simple, sauf que

1. Il faut distinguer les lignes associées au champ ϕ_1 de celles associées au champ ϕ_2 , par exemple en utilisant des lignes tiretées pour le champ ϕ_2 et des lignes pleines pour le champ ϕ_1 .
2. Il y a deux types de vertex, chacun associé à un facteur ig : ceux desquels émergent deux lignes de type 1 et une ligne de type 2, et ceux desquels émergent une ligne de type 1 et deux lignes de type 2.

A Dessinez tous les diagrammes qui contribuent, à l'ordre g^2 , à l'amplitude du processus de collision $1 + 1 \rightarrow 1 + 1$ (c'est-à-dire la collision de deux particules de type 1 donnant encore deux particules de type 1 dans l'état final).

B Dessinez tous les diagrammes qui contribuent, à l'ordre g^2 , à l'amplitude du processus de collision $1 + 1 \rightarrow 2 + 2$ (c'est-à-dire la collision de deux particules de type 1 donnant cette fois deux particules de type 2 dans l'état final).

C Dessinez tous les diagrammes qui contribuent, à l'ordre g^2 , à l'amplitude du processus de collision $1 + 2 \rightarrow 1 + 2$ (c'est-à-dire la collision de deux particules de types différents donnant encore deux particules de types différents dans l'état final).

Problème 2.7 production de particules par une source externe

Un champ scalaire $\phi(\mathbf{r}, t)$ est couplé à une fonction oscillante $f(\mathbf{r})\cos\Omega t$ via un terme qui s'ajoute au à la densité lagrangienne habituelle :

$$\mathcal{L}_1 = g\phi(\mathbf{r}, t)f(\mathbf{r})\cos\Omega t \quad (2.137)$$

A Comment l'équation de Klein-Gordon est-elle modifiée par l'ajout de ce terme?

B Écrivez le hamiltonien correspondant à \mathcal{L}_1 en fonction des opérateurs de création et d'annihilation ($a_{\mathbf{p}}, a_{\mathbf{p}}^\dagger$) et de la transformée de Fourier $f_{\mathbf{p}}$ définie de la manière habituelle :

$$f(\mathbf{r}) = \frac{1}{\mathcal{V}} \sum_{\mathbf{p}} f_{\mathbf{p}} e^{i\mathbf{p}\cdot\mathbf{r}} \quad (2.138)$$

C La règle d'or de Fermi prend la forme suivante pour un système sur lequel s'applique une perturbation dépendant du temps $H_1 = V \cos\Omega t$:

$$\omega_{i \rightarrow f} = 2\pi \frac{1}{4} |M_{fi}|^2 [\delta(E_f - E_i - \Omega) + \delta(E_f - E_i + \Omega)] \quad (2.139)$$

où, au premier ordre, $M_{fi} = \langle f | V | i \rangle$. Comme la perturbation dépend du temps, l'énergie n'est plus conservée, mais peut changer par $\pm\Omega$ si la perturbation est harmonique.

À partir de cette formule, calculez le taux de production Γ de particules causé par la source f . Supposez que f ne dépend que de la distance $|\mathbf{r}|$ par rapport à l'origine.

D Supposons que f soit une fonction sans unités. Quelles sont les unités de g ? Les unités finales du taux de production sont-elles correctes?

CHAPITRE 3

PARTICULES DE SPIN 1/2

A Spin et transformations de Lorentz

L'équation de Klein-Gordon (ou, si on préfère, le champ correspondant) ne peut pas décrire des particules de spin $\frac{1}{2}$, comme les électrons ou autres particules de matière. Lors d'une rotation, par exemple, les deux composantes d'un spineur se combinent entre elles, alors qu'un champ scalaire reste inchangé. Le problème est ici de trouver comment un spineur se transforme quand on passe d'un référentiel à l'autre, de manière à pouvoir écrire une équation d'onde pour un spineur qui soit la même dans tous les référentiels. Ceci nous mènera à l'équation de Dirac. Alors seulement pourrons-nous décrire des électrons relativistes.

3.A.1 Rappel sur les rotations

Considérons une base orthonormée $\{\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3\}$ dans l'espace à trois dimensions. Supposons qu'une deuxième base orthonormée, $\{\mathbf{e}'_1, \mathbf{e}'_2, \mathbf{e}'_3\}$, soit obtenue de la première en effectuant une rotation d'angle θ par rapport à un axe \mathbf{n} . La relation entre les deux bases est alors la suivante :

$$\mathbf{e}'_i = \mathcal{R}_{i1}\mathbf{e}_1 + \mathcal{R}_{i2}\mathbf{e}_2 + \mathcal{R}_{i3}\mathbf{e}_3 = \mathcal{R}_{ij}\mathbf{e}_j \quad (3.1)$$

où la matrice \mathcal{R} est orthogonale : $\mathcal{R}^T \mathcal{R} = 1$. Un vecteur, comme la position \mathbf{r} , peut alors être exprimé dans l'une ou l'autre base :

$$\mathbf{r} = x_i \mathbf{e}_i = x'_i \mathbf{e}'_i = x'_i \mathcal{R}_{ij} \mathbf{e}_j = x'_j \mathcal{R}_{ji} \mathbf{e}_i \quad (3.2)$$

En faisant correspondre les coefficients de \mathbf{e}_i de part et d'autre on trouve

$$x_i = x'_j \mathcal{R}_{ji} \quad \text{ou encore, à l'inverse,} \quad x'_i = \mathcal{R}_{ij} x_j \quad (3.3)$$

Par définition, une quantité à trois composantes (A_1, A_2, A_3) est un *vecteur* si, lors d'une rotation des axes, elle se transforme exactement comme les composantes de la position, soit $A'_i = \mathcal{R}_{ij} A_j$. Précisons qu'il s'agit ici d'une transformation *passive*, c'est-à-dire que ce sont les axes des coordonnées qui sont en rotation et non pas les quantités physiques elles-mêmes. Par contre, on peut également

interpréter la même transformation du point de vue *actif*, c'est-à-dire en supposant que le vecteur \mathbf{A} change et que la base est inchangée.

Par exemple, une rotation d'angle θ par rapport à \mathbf{z} correspond à la matrice

$$\mathcal{R}(\mathbf{z}, \theta) = \begin{pmatrix} \cos \theta & \sin \theta & 0 \\ -\sin \theta & \cos \theta & 0 \\ 0 & 0 & 1 \end{pmatrix} \quad (3.4)$$

Les matrices de rotation $\mathcal{R}(\mathbf{n}, \theta)$ ont un déterminant égal à l'unité. En effet, la condition $\tilde{\mathcal{R}}\mathcal{R} = 1$, lorsqu'on en prend le déterminant, donne $(\det \mathcal{R})^2 = 1$. De plus, pour qu'une rotation puisse être obtenue de manière continue à partir de la transformation identité ($\theta = 0$), il faut que le déterminant soit 1 et non -1 . L'ensemble des matrices orthogonales d'ordre 3 de déterminant unité forment un groupe, noté $SO(3)$ (pour *special orthogonal*). Ce groupe est décrit par trois paramètres réels : deux pour spécifier la direction \mathbf{n} et un pour l'angle de rotation θ .

Rotations infinitésimales Considérons une matrice de rotation $\mathcal{R}(\mathbf{n}, \delta\theta)$ par un angle infinitésimal $\delta\theta$. Cette matrice doit être très proche de l'identité. On peut donc l'écrire, au premier ordre, en $\delta\theta$, comme

$$\mathcal{R}(\mathbf{n}, \delta\theta) \approx 1 + \delta\theta S \quad (3.5)$$

où S est une matrice. La condition $\tilde{\mathcal{R}}\mathcal{R} = 1$ devient alors, toujours au premier ordre en $\delta\theta$,

$$(1 + \delta\theta \tilde{S})(1 + \delta\theta S) = 1 \implies \tilde{S} + S = 0 \quad (3.6)$$

Autrement dit, la matrice S doit être antisymétrique. Comme les matrices réelles antisymétriques sont représentées par trois paramètres réels, on voit que la correspondance est exhaustive. En vertu de l'identité $\text{tr} \ln A = \ln \det A$ pour toute matrice, on constate que la condition $\det \mathcal{R} = 1$ est automatiquement respectée, car les composantes diagonales de S sont nulles et donc $\text{tr} S = 0$. Dans le cas de la matrice (3.4), un développement de Taylor autour de $\theta = 0$ nous montre immédiatement que la matrice S correspondante est

$$S_3 = \begin{pmatrix} 0 & 1 & 0 \\ -1 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} \quad (3.7)$$

En général, pour une rotation infinitésimale autour de l'axe i , la matrice S_i est simplement ϵ_{ijk} . Il tombe sous le sens que la matrice S associée à une rotation autour d'un axe \mathbf{n} général est $S = n_i S_i$:

$$\mathcal{R}(\mathbf{n}, \delta\theta) \approx 1 + \delta\theta n_i S_i \quad (3.8)$$

Une rotation d'angle fini θ peut alors être obtenue en appliquant N fois une rotation d'angle θ/N dans la même direction, ce qui donne, dans la limite $N \rightarrow \infty$,

$$\mathcal{R}(\mathbf{n}, \theta) = \lim_{N \rightarrow \infty} \left(1 + \frac{\theta}{N} n_i S_i \right)^N = e^{\theta n_i S_i} \quad (3.9)$$

d'après la définition de l'exponentielle.

Moment cinétique et rotations Soit une fonction $f(\mathbf{r})$ des coordonnées. Comment une telle fonction est-elle affectée par une rotation infinitésimale *active*? C'est-à-dire par un changement $(x, y, z) \rightarrow (x', y', z')$ infinitésimal de ses arguments? Puisque

$$x'_i = \mathcal{R}_{ij} x_j \approx (\delta_{ij} + \delta\theta \epsilon_{kij} n_k) x_j = x_i + \delta\theta \epsilon_{kij} n_k x_j, \quad (3.10)$$

le changement correspondant de la fonction f est

$$\delta f = f(x', y', z') - f(x, y, z) = \frac{\partial f}{\partial x_i} \delta x_i = \frac{\partial f}{\partial x_i} \delta\theta \epsilon_{kij} n_k x_j = -i \delta\theta \mathbf{L} \cdot \mathbf{n} f \quad (3.11)$$

où nous avons introduit l'opérateur différentiel \mathbf{L} dont les composantes sont

$$L_k = i \epsilon_{kij} x_j \frac{\partial}{\partial x_i} = -i \epsilon_{kij} x_i \frac{\partial}{\partial x_j} = \epsilon_{kij} x_i \hat{p}_j \quad (3.12)$$

où \hat{p}_j est l'opérateur de l'impulsion (composante j). Nous reconnaissons en \mathbf{L} l'opérateur du moment cinétique orbital. Lors d'une rotation finie, la fonction f doit alors se transformer ainsi :

$$f' = f(\mathbf{r}') = e^{-i\theta \mathbf{n} \cdot \mathbf{L}} f(\mathbf{r}) \quad (3.13)$$

3.A.2 Spineurs

La théorie du moment cinétique nous apprend à construire des espaces vectoriels minimaux dans lesquels les composantes du moment cinétique \mathbf{J} ont une action fermée bien définie. L'opérateur \mathbf{J} est une abstraction de l'opérateur \mathbf{L} , c'est-à-dire que ses composantes obéissent aux mêmes relations de commutation

$$[J_i, J_j] = i \epsilon_{ijk} J_k \quad (3.14)$$

mais qu'elles ne s'appliquent pas nécessairement à des fonctions des coordonnées $f(\mathbf{r})$. Les différents espaces vectoriels possibles sont caractérisés par un nombre j qui prend les valeurs entières ou demi-entières positives : $j = 0, \frac{1}{2}, 1, \frac{3}{2}, \dots$. La dimension de l'espace est alors de $2j + 1$. L'action d'une rotation dans cet espace de Hilbert est donnée par l'opérateur

$$R(\mathbf{n}, \theta) = e^{-i\theta \mathbf{n} \cdot \mathbf{J}} \quad (3.15)$$

En particulier, le cas $j = 0$ décrit un état unique invariant par rotation, alors que le cas $j = 1$ décrit un espace à trois états, qui peut contenir en particulier les trois composantes d'un vecteur.

Le cas non trivial le plus simple est $j = \frac{1}{2}$. Dans l'espace de Hilbert correspondant, de dimension 2, les composantes de \mathbf{J} sont représentées par les trois matrices de Pauli :

$$\mathbf{J} = \frac{1}{2} \boldsymbol{\sigma} = \frac{1}{2} (\sigma_1, \sigma_2, \sigma_3) \quad (3.16)$$

où

$$\sigma_1 = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \quad \sigma_2 = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix} \quad \sigma_3 = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \quad (3.17)$$

L'opérateur de rotation correspondant est

$$\exp\left(-i\frac{\theta}{2}\mathbf{n}\cdot\boldsymbol{\sigma}\right) \quad (3.18)$$

Nous allons maintenant définir la matrice

$$R(\mathbf{n}, \theta) = \exp\left(i\frac{\theta}{2}\mathbf{n}\cdot\boldsymbol{\sigma}\right) = \cos\frac{\theta}{2} + i\mathbf{n}\cdot\boldsymbol{\sigma}\sin\frac{\theta}{2} \quad (3.19)$$

où la dernière égalité se démontre simplement en appliquant le développement de Taylor de l'exponentielle, et en réalisant que $(\mathbf{n}\cdot\boldsymbol{\sigma})^2 = 1$.

Un *spineur* par définition, est un objet à deux composantes $\psi = (\psi_1, \psi_2)$ qui, lors d'une rotation, se transforme à l'aide de la matrice (3.19) :

$$\psi' = R(\mathbf{n}, \theta)\psi \quad (3.20)$$

Pourquoi la transformation (3.20), à l'aide de la matrice (3.19), est-elle définie avec un exposant de signe contraire que dans l'opérateur de rotation (3.18)? La raison est que la transformation (3.20) est définie de manière à ce que la combinaison $\psi^\dagger\boldsymbol{\sigma}\psi$ se transforme comme un vecteur, ce qui sera démontré plus loin. C'est aussi une question de différence entre transformation passive et active. Prenons l'exemple d'une translation par un vecteur \mathbf{a} , définie par un opérateur unitaire $T(\mathbf{a})$, donc tel que $T^\dagger(\mathbf{a}) = T^{-1}(\mathbf{a}) = T(-\mathbf{a})$. L'effet de cet opérateur de translation sur les états propres de la position est $T(\mathbf{a})|\mathbf{r}\rangle = |\mathbf{r} + \mathbf{a}\rangle$, au sens actif de la transformation. Par contre, l'effet sur la fonction d'onde $\psi(\mathbf{r}) = \langle\mathbf{r}|\psi\rangle$ est le suivant :

$$\psi(\mathbf{r}) \rightarrow \psi'(\mathbf{r}) = \langle\mathbf{r}|\psi'\rangle = \langle\mathbf{r}|T(\mathbf{a})|\psi\rangle = \langle T^\dagger(\mathbf{a})\mathbf{r}|\psi\rangle = \langle T(-\mathbf{a})\mathbf{r}|\psi\rangle = \langle\mathbf{r} - \mathbf{a}|\psi\rangle = \psi(\mathbf{r} - \mathbf{a}) \quad (3.21)$$

Autrement dit, l'effet de la translation sur l'argument de la fonction d'onde est l'opposé de son effet actif sur la coordonnée. C'est la même chose qui se produit pour le spineur, qui joue le rôle d'une fonction d'onde ici : l'effet de la rotation sur le spineur est l'opposé de son effet sur un état quantique, défini par l'opérateur (3.18)

SU(2) Les matrices R définies par (3.19) ont la propriété que $R^{-1} = R^\dagger$; autrement dit, elles sont unitaires. Cela provient immédiatement du fait que les composantes de \mathbf{J} sont hermitiennes et donc que

$$R^\dagger = \left(e^{i\theta\mathbf{n}\cdot\boldsymbol{\sigma}/2}\right)^\dagger = e^{-i\theta\mathbf{n}\cdot\boldsymbol{\sigma}/2} = R^{-1} \quad (3.22)$$

De plus, le déterminant de R est l'unité : $\det R = 1$ (on dit que R est *unimodulaire*). En effet,

$$\ln \det R = \text{tr} \ln R = i \text{tr} (\mathbf{n}\cdot\boldsymbol{\sigma}/2) = 0 \quad (3.23)$$

car les matrices de Pauli sont sans trace. Les matrices R définies par (3.19) forment en fait un groupe appelé $SU(2)$, formé de toutes les matrices unitaires d'ordre 2 et de déterminant unité (voir l'annexe 7.D à cet effet). Comme trois paramètres sont nécessaires pour paramétrer un élément de $SU(2)$, on voit que la relation entre $SU(2)$ et les rotations est exhaustive.

Spineurs et vecteurs La théorie du moment cinétique nous apprend que le produit tensoriel de deux espaces associés à $j = \frac{1}{2}$ est un espace de dimension 4 qui est la somme directe de deux espaces décrits respectivement par $j = 0$ et $j = 1$. Dit autrement, deux spins $\frac{1}{2}$ peuvent former un singulet ($j = 0$) et un triplet ($j = 1$). En fonction de deux spineurs ψ et χ , cette affirmation revient à dire qu'on peut construire une expression quadratique en ψ et χ qui est invariante par rotation (le singulet), et une autre expression qui se comporte comme un vecteur (le triplet).

Le singulet est évidemment le produit $\psi^\dagger \chi$. Lors d'une rotation, il reste inchangé :

$$(\psi^\dagger \chi)' = \psi^\dagger R^\dagger R \chi = \psi^\dagger \chi \quad \text{car} \quad R^\dagger R = 1 \quad (3.24)$$

Le triplet, lui, est donné par l'expression suivante :

$$\psi^\dagger \boldsymbol{\sigma} \chi \quad (3.25)$$

Pour démontrer que $\psi^\dagger \boldsymbol{\sigma} \chi$ est bel et bien un vecteur, il nous suffit que démontrer que, étant donné un vecteur quelconque \mathbf{A} , l'expression $\psi^\dagger \boldsymbol{\sigma} \cdot \mathbf{A} \chi = \psi^\dagger \sigma_i \chi A_i$ est un scalaire. Or, lors d'une rotation, cette quantité se transforme ainsi :

$$(\psi^\dagger \sigma_i \chi A_i)' = \psi^\dagger R^\dagger \sigma_i R \chi \mathcal{R}_{ij} A_j = \psi^\dagger R^\dagger \sigma_j R \chi \mathcal{R}_{ji} A_i \quad (3.26)$$

où $\mathcal{R}(\mathbf{n}, \theta)$ est la matrice de rotation habituelle et où on a simplement échangé les deux indices muets dans la dernière égalité. Comme cette quantité doit être invariante, donc égale à $\psi^\dagger \sigma_i \chi A_i$, et ce pour tout vecteur \mathbf{A} et pour tous spineurs ψ et χ , nous avons nécessairement la relation

$$R^\dagger \sigma_j R \mathcal{R}_{ji} = \sigma_i \quad (3.27)$$

Comme l'orthogonalité de la matrice \mathcal{R} s'exprime comme $\mathcal{R}_{ji} \mathcal{R}_{ki} = \delta_{jk}$, on peut la mettre à profit en multipliant l'égalité ci-dessus par \mathcal{R}_{ki} pour obtenir

$$R^\dagger \sigma_k R = \mathcal{R}_{ki} \sigma_i \quad (3.28)$$

Cette relation peut aussi être inversée, en remplaçant chaque matrice par son inverse, sachant que $R^{-1} = R^\dagger$ et $\mathcal{R}^{-1} = \tilde{\mathcal{R}}$:

$$R \sigma_k R^\dagger = \mathcal{R}_{ik} \sigma_i \quad (3.29)$$

Comme exemple, considérons une rotation d'angle θ par rapport à \mathbf{z} . La matrice R correspondante est

$$R = \cos \frac{\theta}{2} + i \sigma_3 \sin \frac{\theta}{2} = \begin{pmatrix} e^{i\theta/2} & 0 \\ 0 & e^{-i\theta/2} \end{pmatrix} \quad (3.30)$$

Appliquons la transformation (3.28) sur chacune des matrices de Pauli :

$$\begin{aligned} R^\dagger \sigma^3 R &= \begin{pmatrix} e^{-i\theta/2} & 0 \\ 0 & e^{i\theta/2} \end{pmatrix} \begin{pmatrix} e^{i\theta/2} & 0 \\ 0 & -e^{-i\theta/2} \end{pmatrix} = \sigma^3 \\ R^\dagger \sigma^1 R &= \begin{pmatrix} e^{-i\theta/2} & 0 \\ 0 & e^{i\theta/2} \end{pmatrix} \begin{pmatrix} 0 & e^{-i\theta/2} \\ e^{i\theta/2} & 0 \end{pmatrix} = \begin{pmatrix} 0 & e^{-i\theta} \\ e^{i\theta} & 0 \end{pmatrix} = \sigma^1 \cos \theta + \sigma^2 \sin \theta \end{aligned}$$

$$R^\dagger \sigma^2 R = \begin{pmatrix} e^{-i\theta/2} & 0 \\ 0 & e^{i\theta/2} \end{pmatrix} \begin{pmatrix} 0 & -i e^{-i\theta/2} \\ i e^{i\theta/2} & 0 \end{pmatrix} = \begin{pmatrix} 0 & -i e^{-i\theta} \\ i e^{i\theta} & 0 \end{pmatrix} = \sigma^2 \cos \theta - \sigma^1 \sin \theta \quad (3.31)$$

Ce qui peut encore s'écrire

$$R^\dagger \begin{pmatrix} \sigma_1 \\ \sigma_2 \\ \sigma_3 \end{pmatrix} R = \begin{pmatrix} \cos \theta & \sin \theta & 0 \\ -\sin \theta & \cos \theta & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} \sigma_1 \\ \sigma_2 \\ \sigma_3 \end{pmatrix} = \mathcal{R} \begin{pmatrix} \sigma_1 \\ \sigma_2 \\ \sigma_3 \end{pmatrix} \quad (3.32)$$

où chaque composante des vecteurs en jeu est une matrice d'ordre 2. On vérifie donc la relation (3.28), à savoir que cette matrice R correspond à une rotation d'angle θ par rapport à l'axe z .

En principe, ces propriétés de transformations nous permettent de suggérer des termes supplémentaires dans l'équation de Schrödinger, faisant intervenir le spin mais préservant l'invariance par rotation. Par exemple, les termes suivants :

$$\boldsymbol{\sigma} \cdot \nabla \psi \quad \boldsymbol{\sigma} \cdot (\mathbf{r} \wedge \nabla) \psi \quad \mathbf{B} \cdot \boldsymbol{\sigma} \psi \quad (3.33)$$

Les deux premiers sont des termes dits de « spin-orbite », le troisième une interaction de type Zeeman avec un champ magnétique.

Représentation d'un vecteur par une matrice hermitienne Une autre façon de considérer la transformation (3.28) est d'établir une correspondance entre un vecteur (par exemple, le vecteur position \mathbf{r}) et une matrice X (hermitienne d'ordre 2), de la manière suivante :

$$X = \mathbf{r} \cdot \boldsymbol{\sigma} = x_i \sigma_i = \begin{pmatrix} z & x - iy \\ x + iy & -z \end{pmatrix} \quad (3.34)$$

Lors d'une rotation, $\mathbf{r} \rightarrow \mathbf{r}' = \mathcal{R} \mathbf{r}$ et donc

$$X \rightarrow X' = \mathbf{r}' \cdot \boldsymbol{\sigma} = \mathcal{R}_{ij} x_j \sigma_i = x_j R \sigma_j R^\dagger = R X R^\dagger \quad (3.35)$$

où on a appliqué la relation (3.29). On peut également écrire la relation inverse

$$X = R^\dagger X' R \quad (3.36)$$

Notons en passant que le déterminant de X est invariant par rapport à cette transformation, car

$$\det X = \det R^\dagger \det X' \det R = \det X' \quad (3.37)$$

Or, ce déterminant n'est autre que l'opposée de la norme carrée du vecteur, donc manifestement invariant par rotation :

$$\det X = -(x^2 + y^2 + z^2) \quad (3.38)$$

3.A.3 Spin et invariance de Lorentz

Comment peut-on généraliser la matrice $R(\mathbf{n}, \theta) \in SU(2)$ à une transformation de Lorentz générale au lieu d'une simple rotation? Une façon simple de le faire est d'associer au quadrivecteur position x^μ (ou en fait à tout autre quadrivecteur) la matrice hermitienne suivante :

$$X(x) = \begin{pmatrix} x^0 - x^3 & -x^1 + ix^2 \\ -x^1 - ix^2 & x^0 + x^3 \end{pmatrix} = x_\mu \sigma^\mu \quad (3.39)$$

où on a introduit la notation $\sigma^\mu = (\mathbb{I}, \sigma_1, \sigma_2, \sigma_3)$. La matrice X est la matrice hermitienne 2×2 la plus générale qui soit. Son déterminant est

$$\det X(x) = (x^0)^2 - (x^1)^2 - (x^2)^2 - (x^3)^2 = x_\mu x^\mu \quad (3.40)$$

et est invariant de Lorentz.

Considérons ensuite la transformation matricielle suivante :

$$N^\dagger X' N = X \quad (3.41)$$

où N est une matrice complexe unimodulaire (c'est-à-dire de déterminant unité). Notons que par définition, $X' = x'_\mu \sigma^\mu$ et donc que la matrice N met en correspondance un quadrivecteur x^μ avec un quadrivecteur transformé x'^μ . Comme le déterminant d'un produit de matrices est le produit des déterminants et que $\det N = \det N^\dagger = 1$, on trouve $\det X' = \det X$ et donc

$$(x'^0)^2 - (x'^1)^2 - (x'^2)^2 - (x'^3)^2 = (x^0)^2 - (x^1)^2 - (x^2)^2 - (x^3)^2 \quad (3.42)$$

C'est donc qu'une matrice N préserve les invariants de Lorentz et correspond donc à une transformation de Lorentz bien précise, spécifiée par une matrice de transformation $4 \times 4 \Lambda$:

$$x'^\mu = \Lambda^\mu{}_\nu x^\nu \quad \text{ou} \quad x'_\mu = \Lambda_\mu{}^\nu x_\nu \quad (3.43)$$

où la matrice $\Lambda_\mu{}^\nu$ est l'inverse de la transposée de $\Lambda^\mu{}_\nu$ (c.-à-d. $\Lambda^\mu{}_\alpha \Lambda_\nu{}^\alpha = \delta^\mu_\nu$). Notons qu'il faut $4 \times 2 = 8$ paramètres pour spécifier une matrice complexe 2×2 générale, et que la condition de déterminant unité impose une contrainte complexe qui réduit le nombre de paramètres à 6. Il faut aussi six paramètres pour spécifier une transformation de Lorentz générale (trois composantes pour la vitesse relative des deux référentiels et trois pour spécifier une rotation relative des axes). La transformation (3.41) peut aussi s'écrire

$$N^\dagger x'_\mu \sigma^\mu N = N^\dagger \Lambda_\mu{}^\nu x_\nu \sigma^\mu N = x_\nu \sigma^\nu \quad (3.44)$$

Comme cette relation vaut pour toute valeur de x , on en déduit que σ^μ se comporte comme un quadrivecteur contravariant, soit

$$N^\dagger \Lambda_\mu{}^\nu \sigma^\mu N = \sigma^\nu \quad \text{ou} \quad \boxed{N^\dagger \sigma^\mu N = \Lambda^\mu{}_\nu \sigma^\nu} \quad (3.45)$$

Le raisonnement ci-dessus aurait tout aussi bien pu être mené en utilisant la matrice suivante :

$$\tilde{X}(x) = \begin{pmatrix} x^0 + x^3 & x^1 - ix^2 \\ x^1 + ix^2 & x^0 - x^3 \end{pmatrix} = x_\mu \tilde{\sigma}^\mu \quad (3.46)$$

où $\tilde{\sigma}^\mu = (\mathbb{I}, -\sigma_1, -\sigma_2, -\sigma_3)$. La transformation

$$M^\dagger \tilde{X}' M = \tilde{X} \quad (3.47)$$

aurait alors mené à la propriété

$$M^\dagger \tilde{\sigma}^\mu M = \Lambda^\mu{}_\nu \tilde{\sigma}^\nu \quad (3.48)$$

Les matrices N et M sont reliées ; on montrera plus bas que $M = (N^{-1})^\dagger$.

Rotations Considérons maintenant des exemples précis de transformations. Revenons premièrement aux rotations, plus précisément à la transformation (3.31). Ce cas correspond à $N = R$, où à $M = R$, ce qui est équivalent. Dans ce cas, $R^\dagger = R^{-1}$ (matrice unitaire) et donc $R^\dagger \sigma^0 R = \sigma^0$: la transformation de Lorentz correspondante n'affecte pas la composante temporelle. L'éq. (3.19) donne la forme de la matrice $R = M = N$ correspondant à une rotation générale. Notons que les relations (3.45) et (3.48) sont identiques dans le cas d'une rotation.

Transformation de Lorentz Considérons maintenant un changement de référentiel, caractérisé par la matrice

$$N = P = \begin{pmatrix} e^{-\eta/2} & 0 \\ 0 & e^{\eta/2} \end{pmatrix} \quad (3.49)$$

Montrons que ceci décrit une transformation de Lorentz de rapidité η dans la direction z . En effet,

$$\begin{aligned} P^\dagger \sigma^0 P &= \begin{pmatrix} e^{-\eta} & 0 \\ 0 & e^{\eta} \end{pmatrix} = \begin{pmatrix} \cosh \eta - \sinh \eta & 0 \\ 0 & \cosh \eta + \sinh \eta \end{pmatrix} = \sigma^0 \cosh \eta - \sigma^3 \sinh \eta \\ P^\dagger \sigma^3 P &= \begin{pmatrix} e^{-\eta} & 0 \\ 0 & -e^{\eta} \end{pmatrix} = \begin{pmatrix} \cosh \eta - \sinh \eta & 0 \\ 0 & -\cosh \eta - \sinh \eta \end{pmatrix} = \sigma^3 \cosh \eta - \sigma^0 \sinh \eta \end{aligned} \quad (3.50)$$

alors que

$$P^\dagger \sigma^1 P = \begin{pmatrix} e^{-\eta/2} & 0 \\ 0 & e^{\eta/2} \end{pmatrix} \begin{pmatrix} 0 & e^{\eta/2} \\ e^{-\eta/2} & 0 \end{pmatrix} = \sigma^1 \quad (3.51)$$

et pareillement pour σ^2 . On trouve donc la transformation de Lorentz, exprimée en fonction de la rapidité η (voir éq. (1.6)), où la vitesse relative des deux référentiels est selon l'axe des z .

On vérifie aussi que changer le signe de $\sigma^{1,2,3}$, donc passer de N à M , requiert changer le signe de η , et donc que la même transformation de Lorentz correspond plutôt à $M = P^{-1}$. Enfin, on montre qu'une transformation de Lorentz associée à une vitesse \mathbf{v} générale correspond plutôt à la matrice

$$P(\hat{\mathbf{v}}, \eta) = \cosh \eta/2 - \hat{\mathbf{v}} \cdot \boldsymbol{\sigma} \sinh \eta/2 \quad (3.52)$$

Une transformation générale (changement de référentiel plus rotation) peut toujours s'écrire comme $N = PR$. La matrice M correspondante est alors $M = P^{-1}R = (N^\dagger)^{-1}$. En somme, la relation entre M et N est

$$M = (N^\dagger)^{-1} \quad (3.53)$$

3.A.4 Spineurs droits et gauches : équations de Weyl

La comparaison avec la transformation des spineurs non relativistes (équ. (3.19)) nous porte à poser simplement qu'un spineur se transforme comme $\psi' = N\psi$ ou $\psi' = M\psi$ lors d'une transformation de Lorentz. Pour distinguer ces deux possibilités, on définira un *spineur droit* ψ_R et un *spineur gauche* ψ_L tels que

$$\psi'_R = N\psi_R \quad \text{et} \quad \psi'_L = M\psi_L \quad (3.54)$$

où

$$\begin{aligned} N &= PR & R(\mathbf{n}, \theta) &= \cos \theta/2 + i\mathbf{n} \cdot \boldsymbol{\sigma} \sin \theta/2 \\ M &= P^{-1}R & P(\hat{\mathbf{v}}, \eta) &= \cosh \eta/2 - \hat{\mathbf{v}} \cdot \boldsymbol{\sigma} \sinh \eta/2 \end{aligned} \quad (3.55)$$

Les qualificatifs droit et gauche seront justifiés plus bas. La distinction entre spineurs gauche et droit n'a pas d'importance si on se limite à des rotations, sans changement de référentiel.

Maintenant que nous savons comment les spineurs se transforment lors d'une transformation générale de Lorentz, nous pouvons construire un lagrangien invariant \mathcal{L} impliquant ces spineurs, afin de construire une théorie relativiste des particules de spin $\frac{1}{2}$. Pour cela, nous devons identifier des invariants construits à partir de ces spineurs. Comme démontré plus haut, $N^\dagger \sigma^\mu N = \Lambda^\mu{}_\nu \sigma^\nu$ et $M^\dagger \tilde{\sigma}^\mu M = \Lambda^\mu{}_\nu \tilde{\sigma}^\nu$. Donc les combinaisons

$$\psi_R^\dagger \sigma^\mu \psi_R \quad \text{et} \quad \psi_L^\dagger \tilde{\sigma}^\mu \psi_L \quad (3.56)$$

se transforment comme des quadri-vecteurs. Par exemple,

$$\psi_R'^\dagger \sigma^\mu \psi'_R = \psi_R^\dagger N^\dagger \sigma^\mu N \psi_R = \Lambda^\mu{}_\nu \psi_R^\dagger \sigma^\nu \psi_R \quad (3.57)$$

D'autre part, comme $M = (N^\dagger)^{-1}$, la combinaison $\psi_R^\dagger \psi_L$ et son conjugué complexe $\psi_L^\dagger \psi_R$ sont des invariants de Lorentz :

$$\psi_R'^\dagger \psi'_L = \psi_R^\dagger N^\dagger M \psi_L = \psi_R^\dagger \psi_L \quad (3.58)$$

Construction des équations de Weyl Les expressions suivantes sont donc des termes possibles d'une densité lagrangienne construite à l'aide de spineurs :¹

$$i\psi_L^\dagger \tilde{\sigma}^\mu \partial_\mu \psi_L \quad i\psi_R^\dagger \sigma^\mu \partial_\mu \psi_R \quad (\psi_L^\dagger \psi_R + \psi_R^\dagger \psi_L) \quad i(\psi_L^\dagger \psi_R - \psi_R^\dagger \psi_L) \quad (3.59)$$

1. Nous excluons des termes du type $\partial_\mu \psi_L^\dagger \partial^\mu \psi_R$, parce qu'ils mènent à des dérivées secondes dans l'équation d'onde. L'équation résultante serait tout de même compatible avec la relativité. La raison fondamentale pour laquelle un tel terme est omis est reliée à la notion de renormalisabilité, expliquée à la fin du chapitre suivant. En quelques mots : les théories des champs sont valables dans la limite continue d'une théorie encore inconnue, et les termes comportant des dérivées en trop grand nombre ou d'ordre trop élevé sont négligeables dans ce contexte.

Notons que toutes ces quantités sont réelles, ce qui nécessite la prémultiplication par i dans les deux premiers cas. En effet, l'action S correspondant au premier terme est

$$S = i \int d^4x \psi_L^\dagger \tilde{\sigma}^\mu \partial_\mu \psi_L \quad (3.60)$$

Son conjugué complexe est alors

$$S^* = -i \int d^4x \partial_\mu \psi_L^\dagger \tilde{\sigma}^\mu \psi_L = -i \int d^4x \{ \partial_\mu (\psi_L^\dagger \tilde{\sigma}^\mu \psi_L) - \psi_L^\dagger \tilde{\sigma}^\mu \partial_\mu \psi_L \} \quad (3.61)$$

Le premier terme entre accolades est une dérivée totale et disparaît de l'action, car il se transforme en intégrale sur une hypersurface d'espace-temps à l'infini par utilisation du théorème de Gauss, et la variation du lagrangien est toujours supposée nulle sur la frontière du domaine d'intégration (ici, l'infini). On trouve donc

$$S^* = -i \int d^4x \{ -\psi_L^\dagger \tilde{\sigma}^\mu \partial_\mu \psi_L \} = S \quad (3.62)$$

L'action est donc réelle, grâce au préfacteur i .

Considérons donc une théorie construite à l'aide d'un seul type de spineur, droit ou gauche. Le lagrangien le plus simple, quadratique en ψ , serait

$$\mathcal{L}_L = i \psi_L^\dagger \tilde{\sigma}^\mu \partial_\mu \psi_L \quad \text{ou} \quad \mathcal{L}_R = i \psi_R^\dagger \sigma^\mu \partial_\mu \psi_R \quad (3.63)$$

Une variation $\delta\psi_R$ du champ ψ_R mène à la variation suivante de l'action S_R :

$$\begin{aligned} \delta S_R &= i \int d^4x \{ \delta\psi_R^\dagger \sigma^\mu \partial_\mu \psi_R + \psi_R^\dagger \sigma^\mu \partial_\mu (\delta\psi_R) \} \\ &= i \int d^4x \{ \delta\psi_R^\dagger (\sigma^\mu \partial_\mu \psi_R) - (\partial_\mu \psi_R^\dagger \sigma^\mu) \delta\psi_R \} \\ &= i \int d^4x \{ \delta\psi_R^\dagger (\sigma^\mu \partial_\mu \psi_R) - [\delta\psi_R^\dagger (\sigma^\mu \partial_\mu \psi_R)]^* \} \\ &= \text{Re} \int d^4x \ i \delta\psi_R^\dagger (\sigma^\mu \partial_\mu \psi_R) \end{aligned} \quad (3.64)$$

(nous avons intégré par parties le deuxième terme). Demander que toute variation réelle (par opposition à complexe) de l'action soit nulle mène à l'équation d'onde suivante :

$$\sigma^\mu \partial_\mu \psi_R = 0 \quad (3.65)$$

On obtient le même résultat en supposant que les variations $\delta\psi_R$ et $\delta\psi_R^\dagger$ sont indépendantes (car ce sont des nombres complexes), ce qui mène directement, à partir de la deuxième des équations ci-haut, aux équations équivalentes

$$\sigma^\mu \partial_\mu \psi_R = 0 \quad \text{et} \quad \partial_\mu \psi_R^\dagger \sigma^\mu = 0 \quad (3.66)$$

(la deuxième équation est la conjuguée de la première).

On montre pareillement qu'une variation de ψ_L dans S_L mène à

$$\tilde{\sigma}^\mu \partial_\mu \psi_L = 0 \quad \text{et} \quad \partial_\mu \psi_L^\dagger \tilde{\sigma}^\mu = 0 \quad (3.67)$$

Ce sont les *équations de Weyl*, et les spineurs complexes à deux composantes sont souvent appelés *spineurs de Weyl*.

Ondes planes Pour voir quel type d'objet est décrit par ces équations, posons une solution de type onde plane :

$$\psi_L = u_L e^{-ip^\mu x_\mu} \quad \psi_R = u_R e^{-ip^\mu x_\mu} \quad (3.68)$$

En substituant dans les équations d'ondes ci-haut, on trouve

$$\tilde{\sigma}^\mu p_\mu u_L = 0 \quad \sigma^\mu p_\mu u_R = 0 \quad (3.69)$$

Supposons de plus que ces ondes se déplacent dans la direction $+z$, de sorte que $p_\mu = (E, 0, 0, -p)$ (composantes covariantes), où $p = |\mathbf{p}| > 0$. Ces équations deviennent explicitement

$$\begin{pmatrix} E+p & 0 \\ 0 & E-p \end{pmatrix} u_L = 0 \quad \begin{pmatrix} E-p & 0 \\ 0 & E+p \end{pmatrix} u_R = 0 \quad (3.70)$$

En supposant que l'énergie est nécessairement positive, on est forcé d'admettre uniquement les solutions suivantes :

$$u_L = \begin{pmatrix} 0 \\ 1 \end{pmatrix} \quad (E = |\mathbf{p}|) \quad \text{et} \quad u_R = \begin{pmatrix} 1 \\ 0 \end{pmatrix} \quad (E = |\mathbf{p}|) \quad (3.71)$$

On constate que les objets décrits par ces solutions ont une masse nulle, car $E = |\mathbf{p}|$. D'autre part, la projection de spin associée à ψ_L est négative, et celle associée à ψ_R est positive. Il existe aussi des solutions à énergie négative : $E = -|\mathbf{p}|$. Dans ces solutions, les projections de spin sont inversées. Nous remettons à plus tard, dans le contexte de l'équation de Dirac, l'étude et l'interprétation de ces solutions. Nous verrons qu'elles décrivent en fait des antiparticules.

Pour conclure, un champ de spineur ψ_R ou ψ_L ne peut décrire, en soi, que des particules de spin $\frac{1}{2}$ sans masse comme les neutrinos (ou les antineutrinos). Une description des particules massives de spin $\frac{1}{2}$, comme l'électron, doit combiner des spineurs droits et gauche, comme nous le verrons plus bas.

3.A.5 Hélicité

On définit l'*hélicité* comme la composante du spin \mathbf{s} le long de la quantité de mouvement :

$$h = \mathbf{s} \cdot \hat{\mathbf{p}} \quad (3.72)$$

L'hélicité d'une particule sans masse est un invariant. En effet, la combinaison $\mathbf{s} \cdot \hat{\mathbf{p}}$ est un produit scalaire, et donc déjà invariant par rotation dans l'espace. Il suffit de montrer ici qu'elle est aussi

invariante par rapport aux transformations de Lorentz dans la direction du mouvement de la particule. Posons justement que la particule se dirige le long de l'axe des z , de sorte que $\hat{\mathbf{p}} = \mathbf{z}$. Dans ce cas, $h = s_z$. Mais lors d'une transformation de Lorentz dans la direction z , le moment cinétique dans cette direction reste inchangé, car en général son expression est $L_z = x p_y - y p_x$ et les composantes de \mathbf{r} et de \mathbf{p} perpendiculaires à la direction de la transformation ne sont pas affectées par la transformation. Cette propriété du moment cinétique s'étend aussi au moment cinétique intrinsèque (spin), même si celui-ci n'est pas exprimé en fonction de la position et de la quantité de mouvement. La seule manière de changer l'hélicité serait de changer la direction de $\hat{\mathbf{p}}$, ce qui peut se faire si la transformation de Lorentz nous permet de « dépasser » la particule, de sorte que $h = \mathbf{s} \cdot \hat{\mathbf{p}}$ change de signe. Mais une particule sans masse se déplace toujours à la vitesse de la lumière et ne peut être dépassée. Donc l'hélicité h est un invariant de Lorentz.

Les solutions trouvées ci-haut montrent que le spineur ψ_L a une hélicité négative, soit un spin antiparallèle à sa vitesse. On dit que ceci correspond à une polarisation circulaire gauche, d'où le qualificatif *gauche* associé à ce type de spineur. Le contraire prévaut pour le spineur ψ_R .

B Équation de Dirac

Un champ de spineur ψ_R ou ψ_L ne peut décrire, en soi, que des particules de spin $\frac{1}{2}$ sans masse comme les neutrinos. Pour décrire une particule massive comme l'électron, nous devons combiner des spineurs droit et gauche. Nous devons écrire une densité lagrangienne invariante de Lorentz impliquant à la fois ψ_R et ψ_L , à l'aide des termes figurant dans (3.59). Si nous demandons en plus que la théorie soit invariante par inversion de l'espace (transformation de parité), il faut aussi que la densité lagrangienne reste la même lorsqu'on intervertit ψ_L et ψ_R . En effet, une inversion de l'espace change la polarisation droite en polarisation gauche et vice-versa. Une autre façon de le voir est que l'inversion de l'espace change le signe du vecteur vitesse (ou quantité de mouvement) mais préserve l'orientation du vecteur moment cinétique, de sorte que l'hélicité $h = \mathbf{s} \cdot \hat{\mathbf{p}}$ change de signe. La seule possibilité simple pour la densité lagrangienne est

$$\mathcal{L}_D = i\psi_L^\dagger \tilde{\sigma}^\mu \partial_\mu \psi_L + i\psi_R^\dagger \sigma^\mu \partial_\mu \psi_R - m(\psi_L^\dagger \psi_R + \psi_R^\dagger \psi_L) \quad (3.73)$$

Nous verrons que le paramètre m correspond à la masse des particules.

Plutôt que de travailler avec des spineurs droit et gauche séparés, on les combine généralement en un seul spineur à quatre composantes, dit *spineur de Dirac* :

$$\psi = \begin{pmatrix} \psi_L \\ \psi_R \end{pmatrix} \quad (3.74)$$

On peut alors écrire la densité lagrangienne ainsi :

$$\mathcal{L} = i\psi^\dagger \begin{pmatrix} \tilde{\sigma}^\mu & 0 \\ 0 & \sigma^\mu \end{pmatrix} \partial_\mu \psi - m\psi^\dagger \begin{pmatrix} 0 & \mathbb{I} \\ \mathbb{I} & 0 \end{pmatrix} \psi \quad (3.75)$$

On définit en outre les matrices 4×4 suivantes :

$$\gamma^0 = \begin{pmatrix} 0 & \mathbb{I} \\ \mathbb{I} & 0 \end{pmatrix} \quad \gamma^k = \begin{pmatrix} 0 & \sigma_k \\ -\sigma_k & 0 \end{pmatrix} \quad (k = 1, 2, 3) \quad (3.76)$$

ce qui équivaut à

$$\gamma^\mu = \gamma^0 \begin{pmatrix} \tilde{\sigma}^\mu & 0 \\ 0 & \sigma^\mu \end{pmatrix} \quad (\mu = 0, 1, 2, 3) \quad (3.77)$$

On les appelle *matrices de Dirac*. Enfin, on définit le *spineur conjugué* $\bar{\psi}$ comme

$$\bar{\psi} = \psi^\dagger \gamma^0 \quad (3.78)$$

La densité lagrangienne peut alors s'écrire de la manière suivante :

$$\mathcal{L}_D = i \bar{\psi} \gamma^\mu \partial_\mu \psi - m \bar{\psi} \psi \quad (3.79)$$

L'avantage de définir le spineur conjugué $\bar{\psi}$ est que la combinaison $\bar{\psi} \psi$ est un invariant de Lorentz, alors que la combinaison $\bar{\psi} \gamma^\mu \psi$ est un quadrivecteur :

$$\bar{\psi} \psi = \psi^\dagger \gamma^0 \psi = \begin{pmatrix} \psi_L^\dagger & \psi_R^\dagger \end{pmatrix} \begin{pmatrix} 0 & \mathbb{I} \\ \mathbb{I} & 0 \end{pmatrix} \begin{pmatrix} \psi_L \\ \psi_R \end{pmatrix} = \psi_L^\dagger \psi_R + \psi_R^\dagger \psi_L \quad (3.80)$$

$$\bar{\psi} \gamma^\mu \psi = \psi^\dagger \gamma^0 \gamma^\mu \psi = \begin{pmatrix} \psi_L^\dagger & \psi_R^\dagger \end{pmatrix} \begin{pmatrix} \tilde{\sigma}^\mu & 0 \\ 0 & \sigma^\mu \end{pmatrix} \begin{pmatrix} \psi_L \\ \psi_R \end{pmatrix} = \psi_L^\dagger \tilde{\sigma}^\mu \psi_L + \psi_R^\dagger \sigma^\mu \psi_R \quad (3.81)$$

Une fois la forme de la densité lagrangienne établie, nous pouvons appliquer le principe de la moindre action et trouver les équations de Lagrange correspondantes. Lors d'une variation $\delta \psi$ du spineur de Dirac, le spineur conjugué varie comme $\delta \bar{\psi} = \delta \psi^\dagger \gamma^0$. La variation de l'action de Dirac qui s'ensuit est

$$\begin{aligned} \delta S[\psi] &= \int d^4x \left\{ \delta \bar{\psi} (i \gamma^\mu \partial_\mu \psi - m \psi) + (i \bar{\psi} \gamma^\mu \partial_\mu \delta \psi - m \bar{\psi} \delta \psi) \right\} \\ &= \int d^4x \left\{ \delta \bar{\psi} (i \gamma^\mu \partial_\mu \psi - m \psi) + (-i \partial_\mu \bar{\psi} \gamma^\mu - m \bar{\psi}) \delta \psi \right\} \end{aligned} \quad (3.82)$$

où on a intégré par parties dans la dernière parenthèse. Comme le champ ψ est complexe, on peut considérer que ses parties réelle et imaginaire varient de manières indépendantes, ce qui se traduit par des variations $\delta \psi$ et $\delta \bar{\psi}$ indépendantes. Le principe de la moindre action se traduit donc par les équations séparées

$$i \gamma^\mu \partial_\mu \psi - m \psi = 0 \quad \text{et} \quad -i \partial_\mu \bar{\psi} \gamma^\mu - m \bar{\psi} = 0 \quad (3.83)$$

On montre que ces équations sont en fait équivalentes, la deuxième étant la conjuguée de la première. Il faut pour cela utiliser la propriété

$$(\gamma^\mu)^\dagger = \gamma^0 \gamma^\mu \gamma^0 \quad (3.84)$$

qu'on démontre trivialement à partir de la définition de γ^μ . La première des équations (3.83) est la célèbre *équation de Dirac*.

3.B.1 Propriétés des matrices de Dirac

Avant d'étudier les solutions à l'équation de Dirac, signalons que les matrices de Dirac ne prennent pas toujours la forme (3.76). Il est possible d'appliquer une transformation unitaire $\psi \rightarrow U\psi$ sur les spineurs de Dirac et de modifier en même temps les matrices de Dirac par $\gamma^\mu \rightarrow U\gamma^\mu U^\dagger$ sans modifier ni l'action, ni l'équation de Dirac.

La représentation (3.76) est qualifiée de *chirale*, alors que nous allons surtout utiliser la *représentation de Dirac* :

$$\gamma^0 = \begin{pmatrix} \mathbb{I} & 0 \\ 0 & -\mathbb{I} \end{pmatrix} \quad \gamma^i = \begin{pmatrix} 0 & \sigma_i \\ -\sigma_i & 0 \end{pmatrix} \quad (3.85)$$

On vérifie que cette représentation s'obtient de (3.76) par la transformation

$$\gamma^\mu \rightarrow U\gamma^\mu U^\dagger \quad \text{où} \quad U = \frac{1}{\sqrt{2}} \begin{pmatrix} \mathbb{I} & \mathbb{I} \\ -\mathbb{I} & \mathbb{I} \end{pmatrix} \quad (3.86)$$

ceci correspond à l'utilisation du spineur

$$\psi = \frac{1}{\sqrt{2}} \begin{pmatrix} \psi_L + \psi_R \\ -\psi_L + \psi_R \end{pmatrix} \quad (3.87)$$

Quelle que soit la représentation, on vérifie que les matrices de Dirac ont les propriétés suivantes :

$$\gamma^\mu \gamma^\nu + \gamma^\nu \gamma^\mu = 2g^{\mu\nu} \quad \gamma^0 \gamma^\mu \gamma^0 = (\gamma^\mu)^\dagger \quad (3.88)$$

3.B.2 Ondes planes

Trouvons maintenant les solutions à l'équation de Dirac sous la forme d'ondes planes :

$$\psi = \psi_{\mathbf{p}} e^{i(\mathbf{p} \cdot \mathbf{r} - Et)} = \psi_{\mathbf{p}} e^{-ip_\mu x^\mu} \quad (3.89)$$

En substituant dans l'équation de Dirac, on trouve

$$(p_\mu \gamma^\mu - m)\psi_{\mathbf{p}} = 0 \quad (3.90)$$

ou encore

$$\left(\gamma^0 E - \sum_j \gamma^j p_j - m \right) \psi_{\mathbf{p}} = 0 \quad (3.91)$$

Après prémultiplication par γ^0 , on trouve

$$\left(E - \sum_j \gamma^0 \gamma^j p_j - m \gamma^0 \right) \psi_{\mathbf{p}} = 0 \quad \text{ou} \quad \begin{pmatrix} m & \mathbf{p} \cdot \boldsymbol{\sigma} \\ \mathbf{p} \cdot \boldsymbol{\sigma} & -m \end{pmatrix} \psi_{\mathbf{p}} = E \psi_{\mathbf{p}} \quad (3.92)$$

(avec la représentation de Dirac (3.85)). Il s'agit d'une équation aux valeurs propres : les valeurs possibles de E sont les valeurs propres de la matrice qui multiplie $\psi_{\mathbf{p}}$ et les spineurs $\psi_{\mathbf{p}}$ sont les vecteurs propres associés. Plus généralement, en multipliant l'équation de Dirac par γ^0 , on trouve une forme qui nous rapproche de l'équation de Schrödinger :

$$i \frac{\partial \psi}{\partial t} = \hat{\mathcal{H}} \psi \quad \text{où} \quad \hat{\mathcal{H}} := -i \sum_j \gamma^0 \gamma^j \frac{\partial}{\partial x^j} + m \gamma^0 = \begin{pmatrix} m & -i \boldsymbol{\sigma} \cdot \nabla \\ -i \boldsymbol{\sigma} \cdot \nabla & -m \end{pmatrix} \quad (3.93)$$

L'équation (3.92) peut donc s'exprimer sous la forme $\hat{\mathcal{H}} \psi_{\mathbf{p}} = E \psi_{\mathbf{p}}$, l'opérateur différentiel $\hat{\mathcal{H}}$ étant le hamiltonien à une particule associé à l'équation de Dirac.

Les valeurs propres se trouvent en résolvant l'équation caractéristique :

$$\begin{vmatrix} m - E & \mathbf{p} \cdot \boldsymbol{\sigma} \\ \mathbf{p} \cdot \boldsymbol{\sigma} & -m - E \end{vmatrix} = (E^2 - m^2 - \mathbf{p}^2)^2 = 0 \quad (3.94)$$

Si on définit $E_{\mathbf{p}} = \sqrt{\mathbf{p}^2 + m^2}$, alors les valeurs propres sont $E = \pm E_{\mathbf{p}}$. Il y a donc des solutions à énergie positive et d'autres à énergie négative.

Accès rapide aux vecteurs propres

Comme le déterminant associé à l'équation caractéristique est relativement incommode à évaluer, décrivons une façon alternative de trouver les valeurs propres. Remarquons à cet effet la relation suivante :

$$(p_\mu \gamma^\mu + m)(p_\nu \gamma^\nu - m) = p^2 - m^2. \quad (3.95)$$

En effet,

$$(p_\mu \gamma^\mu + m)(p_\nu \gamma^\nu - m) = p_\mu p_\nu \gamma^\mu \gamma^\nu - m^2 \quad (3.96)$$

et

$$p_\mu p_\nu \gamma^\mu \gamma^\nu = \frac{1}{2} p_\mu p_\nu (\gamma^\mu \gamma^\nu + \gamma^\nu \gamma^\mu) = p_\mu p_\nu g^{\mu\nu} = p^2. \quad (3.97)$$

Cette relation est tout aussi vraie dans l'ordre inverse :

$$(p_\mu \gamma^\mu - m)(p_\nu \gamma^\nu + m) = p^2 - m^2. \quad (3.98)$$

Donc, en appliquant par la gauche $(p_\nu \gamma^\nu + m)$ sur l'équation de Dirac $(p_\mu \gamma^\mu - m) \psi_{\mathbf{p}} = 0$, on trouve

$$(p_\mu \gamma^\mu + m)(p_\nu \gamma^\nu - m) \psi_{\mathbf{p}} = (p^2 - m^2) \psi_{\mathbf{p}} = 0 \quad (3.99)$$

Ce qui entraîne nécessairement que $p^2 = m^2$, ou encore $E = \pm E_{\mathbf{p}}$.

Déterminons maintenant la forme des vecteurs propres associés. Pour faciliter les choses, commençons par le cas de particules immobiles ($\mathbf{p} = 0$). L'équation aux valeurs propres prend alors la forme très simple

$$\begin{pmatrix} m & 0 \\ 0 & -m \end{pmatrix} \psi_0 = E \psi_0 \quad (3.100)$$

et les vecteurs propres normalisés, notés u_i ($i = 1, 2, 3, 4$) sont simplement

$$E = m : \quad u_1 = \begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \end{pmatrix} \quad u_2 = \begin{pmatrix} 0 \\ 1 \\ 0 \\ 0 \end{pmatrix} \quad E = -m : \quad u_3 = \begin{pmatrix} 0 \\ 0 \\ 1 \\ 0 \end{pmatrix} \quad u_4 = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 1 \end{pmatrix} \quad (3.101)$$

Utilisons ensuite un truc pour trouver les solutions associées à des valeurs non nulles de \mathbf{p} . Nous devons résoudre $(p_\mu \gamma^\mu - m)\psi_{\mathbf{p}} = 0$. Or, tout spineur de la forme $(p_\mu \gamma^\mu + m)\psi_0$ est une solution à cette équation, en vertu de l'identité (3.98). Il suffit donc d'appliquer l'expression matricielle $(p_\nu \gamma^\nu + m)$ aux solutions trouvées ci-haut pour $\mathbf{p} = 0$ pour obtenir les solutions correspondant à une quantité de mouvement $\mathbf{p} \neq 0$.

Notons premièrement que

$$p_\nu \gamma^\nu + m = \begin{pmatrix} E_{\mathbf{p}} + m & -\mathbf{p} \cdot \boldsymbol{\sigma} \\ \mathbf{p} \cdot \boldsymbol{\sigma} & -E_{\mathbf{p}} + m \end{pmatrix} = \begin{pmatrix} p^0 + m & 0 & -p_z & -p_x + i p_y \\ 0 & p^0 + m & -p_x - i p_y & p_z \\ p_z & p_x - i p_y & -p^0 + m & 0 \\ p_x + i p_y & -p_z & 0 & -p^0 + m \end{pmatrix} \quad (3.102)$$

Les solutions à énergie positive ($p^0 = E_{\mathbf{p}}$) sont donc :

$$(p_\nu \gamma^\nu + m)u_1 = \begin{pmatrix} E_{\mathbf{p}} + m \\ 0 \\ p_z \\ p_x + i p_y \end{pmatrix} \quad \text{et} \quad (p_\nu \gamma^\nu + m)u_2 = \begin{pmatrix} 0 \\ E_{\mathbf{p}} + m \\ p_x - i p_y \\ -p_z \end{pmatrix} \quad (3.103)$$

Normalisation des solutions Ces spineurs ne sont cependant pas normalisés. Il est préférable d'utiliser des spineurs normalisés pour effectuer un développement du champ de Dirac sur ses modes propres. La norme au carré des spineurs ci-haut est :

$$(E_{\mathbf{p}} + m)^2 + \mathbf{p}^2 = (E_{\mathbf{p}} + m)^2 + E_{\mathbf{p}}^2 - m^2 = 2E_{\mathbf{p}}(m + E_{\mathbf{p}}) \quad (3.104)$$

En divisant par cette norme (la racine carrée de l'expression ci-dessus), on trouve des spineurs normalisés, que nous noterons $u_{\mathbf{p},\alpha}$:

$$u_{\mathbf{p},1} = \frac{1}{\sqrt{2E_{\mathbf{p}}(E_{\mathbf{p}} + m)}} \begin{pmatrix} E_{\mathbf{p}} + m \\ 0 \\ p_z \\ p_x + i p_y \end{pmatrix} \quad u_{\mathbf{p},2} = \frac{1}{\sqrt{2E_{\mathbf{p}}(E_{\mathbf{p}} + m)}} \begin{pmatrix} 0 \\ E_{\mathbf{p}} + m \\ p_x - i p_y \\ -p_z \end{pmatrix} \quad (3.105)$$

Solutions à énergie négative Pour les solutions à énergie négative, la même procédure donne ($p^0 = -E_{\mathbf{p}}$) :

$$(p_\nu \gamma^\nu + m)u_3 = \begin{pmatrix} -p_z \\ -p_x - ip_y \\ E_{\mathbf{p}} + m \\ 0 \end{pmatrix} \quad \text{et} \quad (p_\nu \gamma^\nu + m)u_4 = \begin{pmatrix} -p_x + ip_y \\ p_z \\ 0 \\ E_{\mathbf{p}} + m \end{pmatrix} \quad (3.106)$$

ce qui devient, après normalisation :

$$u_{\mathbf{p},3} = \frac{1}{\sqrt{2E_{\mathbf{p}}(E_{\mathbf{p}} + m)}} \begin{pmatrix} -p_z \\ -p_x - ip_y \\ E_{\mathbf{p}} + m \\ 0 \end{pmatrix} \quad u_{\mathbf{p},4} = \frac{1}{\sqrt{2E_{\mathbf{p}}(E_{\mathbf{p}} + m)}} \begin{pmatrix} -p_x + ip_y \\ p_z \\ 0 \\ E_{\mathbf{p}} + m \end{pmatrix} \quad (3.107)$$

Ces quatre « spineurs propres » sont orthonormés :

$$u_{\mathbf{p},s}^\dagger u_{\mathbf{p},s'} = \delta_{ss'} \quad (s, s' = 1, 2, 3, 4) \quad (3.108)$$

Il est clair que les solutions trouvées correspondent à des particules de spin $\frac{1}{2}$, mais l'existence de quatre solutions (au lieu de deux) et la présence de solutions à énergie négative soulèvent des questions ! Nous y répondrons dans la section suivante, en procédant à la deuxième quantification de l'équation de Dirac.

C Quantification du champ de Dirac

3.C.1 Hamiltonien et relations de commutation

Appliquons au champ de Dirac la définition générale (2.23) du hamiltonien d'un champ :

$$\mathcal{H} = \sum_{\alpha} \pi_{\alpha} \dot{\psi}_{\alpha} - \mathcal{L} \quad \text{où} \quad \pi_{\alpha} = \frac{\partial \mathcal{L}}{\partial \dot{\psi}_{\alpha}} \quad (3.109)$$

D'après l'expression (3.79) du lagrangien, on trouve

$$\pi_{\alpha} = i(\bar{\psi} \gamma^0)_{\alpha} = i\psi_{\alpha}^* \quad \pi_{\alpha}^* = -i\psi_{\alpha} \quad (3.110)$$

Les règles de quantification canonique mènent alors aux relations de commutation suivantes :

$$[\psi_{\alpha}(\mathbf{r}), \pi_{\beta}(\mathbf{r}')] = i\delta_{\alpha\beta} \delta(\mathbf{r} - \mathbf{r}') \quad [\psi_{\alpha}(\mathbf{r}), \pi_{\beta}^{\dagger}(\mathbf{r}')] = 0 \quad (3.111)$$

car les variables dynamiques conjuguées complexes sont considérées indépendantes. Notons que la variable conjuguée complexe devient, après quantification, l'opérateur conjugué hermitien. Cela revient à écrire

$$[\psi_\alpha(\mathbf{r}), \psi_\beta^\dagger(\mathbf{r}')] = \delta_{\alpha\beta} \delta(\mathbf{r} - \mathbf{r}') \quad [\psi_\alpha(\mathbf{r}), \psi_\beta(\mathbf{r}')] = 0 \quad (3.112)$$

Afin de définir l'équivalent pour le champ de Dirac des opérateurs de création et d'annihilation du champ scalaire, décomposons le champ de Dirac sur une base de solutions à l'équation de Dirac :

$$\psi(\mathbf{r}) = \frac{1}{\sqrt{\mathcal{V}}} \sum_{\mathbf{p}} \sum_{s=1}^4 c_{\mathbf{p},s} u_{\mathbf{p},s} e^{i\mathbf{p}\cdot\mathbf{r}} \quad (3.113)$$

Cette série peut être considérée comme une définition de nouveaux opérateurs $c_{\mathbf{p},s}$. Il est plus pratique d'avoir la relation inverse de (3.113), ce qui est simple, étant donné que les solutions figurant dans le développement (3.113) sont orthonormées :

$$c_{\mathbf{p},s} = \frac{1}{\sqrt{\mathcal{V}}} \int d^3r u_{\mathbf{p},s,\alpha}^* \psi_\alpha(\mathbf{r}) e^{-i\mathbf{p}\cdot\mathbf{r}} \quad (3.114)$$

$$c_{\mathbf{p},s}^\dagger = \frac{1}{\sqrt{\mathcal{V}}} \int d^3r u_{\mathbf{p},s,\alpha} \psi_\alpha^\dagger(\mathbf{r}) e^{i\mathbf{p}\cdot\mathbf{r}} \quad (3.115)$$

On peut alors calculer le commutateur suivant :

$$[c_{\mathbf{p},s}, c_{\mathbf{p}',s'}^\dagger] = \frac{1}{\mathcal{V}} \int d^3r d^3r' u_{\mathbf{p},s,\alpha}^* u_{\mathbf{p}',s',\beta} e^{i(\mathbf{p}'\cdot\mathbf{r}' - \mathbf{p}\cdot\mathbf{r})} [\psi_\alpha(\mathbf{r}), \psi_\beta^\dagger(\mathbf{r}')] \quad (3.116)$$

Les relations de commutation (3.112) nous permettent de simplifier l'expression ci-dessus :

$$[c_{\mathbf{p},s}, c_{\mathbf{p}',s'}^\dagger] = \frac{1}{\mathcal{V}} \int d^3r u_{\mathbf{p},s,\alpha}^* u_{\mathbf{p}',s',\beta} \delta_{\alpha\beta} e^{i\mathbf{r}\cdot(\mathbf{p}' - \mathbf{p})} \quad (3.117)$$

En utilisant les relations d'orthogonalité (3.108) et en calculant l'intégrale, on trouve

$$[c_{\mathbf{p},s}, c_{\mathbf{p}',s'}^\dagger] = \delta_{\mathbf{p},\mathbf{p}'} \delta_{s,s'} \quad (3.118)$$

De même, on déduit, plus facilement, que

$$[c_{\mathbf{p},s}, c_{\mathbf{p}',s'}] = 0 \quad (3.119)$$

Ainsi, les règles de quantification canonique mènent à des relations de commutation pour $c_{\mathbf{p},s}$ et $c_{\mathbf{p},s}^\dagger$ qui sont les mêmes que pour les opérateurs de création et d'annihilation du champ scalaire. On pourrait donc, logiquement, appliquer les mêmes raisonnements que précédemment et définir les états propres

$$|0\rangle, \quad c_{\mathbf{p},s}^\dagger |0\rangle, \quad c_{\mathbf{p},s}^\dagger c_{\mathbf{p}',s'}^\dagger |0\rangle, \text{ etc.} \quad (3.120)$$

3.C.2 Fermions et anticommutateurs

Les relations de commutation (3.118), quoiqu'elles dérivent des règles de quantification habituelles, ne peuvent s'appliquer aux fermions. La raison en est que les états à plusieurs particules dans (3.120) sont symétriques lorsqu'on permute les opérateurs de création, ce qui est la marque des bosons. Pour que les particules soient des fermions, il faudrait plutôt que les états à plusieurs particules soient antisymétriques.

La façon d'y arriver, qui peut sembler tout à fait *ad hoc* dans le présent contexte, est de remplacer tous les commutateurs par des *anticommutateurs*. On définit l'anticommutateur $\{A, B\}$ de deux opérateurs A et B comme

$$\{A, B\} := AB + BA \quad (3.121)$$

Certains auteurs utilisent plutôt la notation $[A, B]_+$. Nous allons donc imposer les relations suivantes au champ de Dirac :

$$\boxed{\{\psi_\alpha(\mathbf{r}), \psi_\beta^\dagger(\mathbf{r}')\} = \delta_{\alpha\beta} \delta(\mathbf{r} - \mathbf{r}')} \quad \boxed{\{\psi_\alpha(\mathbf{r}), \psi_\beta(\mathbf{r}')\} = 0} \quad (3.122)$$

En répétant les calculs de la section précédente, on trouve alors

$$\boxed{\{c_{\mathbf{p},s}, c_{\mathbf{p}',s'}^\dagger\} = \delta_{\mathbf{p},\mathbf{p}'} \delta_{ss'}} \quad \boxed{\{c_{\mathbf{p},s}, c_{\mathbf{p}',s'}\} = 0} \quad (3.123)$$

Oscillateur fermionique Considérons un seul mode (\mathbf{p}, s) , et désignons provisoirement l'opérateur d'annihilation correspondant par c . On a donc les relations suivantes :

$$\{c, c^\dagger\} = 1 \quad \{c, c\} = 0 \quad (3.124)$$

La deuxième de ces relations est équivalente à $c^2 = 0$: l'opérateur d'annihilation est nilpotent, de même, naturellement, que son conjugué hermitien : $(c^\dagger)^2 = 0$. Si on définit l'état vide $|0\rangle$ par la relation $c|0\rangle = 0$, alors il n'y a qu'un seul autre état dans l'espace de Hilbert : l'état $|1\rangle := c^\dagger|0\rangle$. En effet, l'application de c^\dagger sur $|1\rangle$ donne zéro car $(c^\dagger)^2 = 0$, et l'application de c sur $|1\rangle$ nous redonne l'état $|0\rangle$:

$$c|1\rangle = c c^\dagger|0\rangle = (c c^\dagger + c^\dagger c)|0\rangle = |0\rangle \quad (3.125)$$

(le deuxième terme entre parenthèses ne contribue pas, car $c|0\rangle = 0$ par construction, et la parenthèse n'est autre que l'anticommutateur).

Les opérateurs c et c^\dagger sont donc réalisés dans un système à deux niveaux, qu'on appelle parfois «oscillateur fermionique», pour accentuer la ressemblance formelle avec un oscillateur harmonique, qui décrit un mode individuel du champ scalaire, ou d'un boson en général. Les relations d'anticommutation se trouvent à imposer naturellement le *principe d'exclusion de Pauli* : il est impossible de construire un état à plusieurs particules contenant plus d'un quantum du même mode, en raison de la relation $(c_{\mathbf{p},s}^\dagger)^2 = 0$.

Notons que l'opérateur $N = c^\dagger c$ a encore l'interprétation du nombre de quanta : ses valeurs propres sont 0 et 1, comme on peut le voir explicitement :

$$N|0\rangle = c^\dagger c|0\rangle = 0 \quad N|1\rangle = c^\dagger c|1\rangle = c^\dagger|0\rangle = |1\rangle \quad (3.126)$$

Remarques :

- ◆ Les anticommutateurs surviennent naturellement si on impose que la variable classique $\psi_\alpha(\mathbf{r})$, avant quantification, ne soit pas un nombre complexe ordinaire, mais plutôt un nombre anticommutatif, ce qu'on appelle plus précisément une variable de Grassmann.
- ◆ L'imposition d'anticommutateurs au lieu de commutateurs peut aussi être dérivée de manière totalement différente, par construction pour ainsi dire, en construisant le champ de Dirac à partir de l'espace des états, ce qui est le contraire de la voie suivie ici.

3.C.3 Hamiltonien de Dirac

Retournons à l'éq. (3.109). Le hamiltonien, après substitution du moment conjugué $\pi_\alpha = i\psi_\alpha^\dagger$, est

$$\begin{aligned} H &= \int d^3r [\pi_\alpha \dot{\psi}_\alpha - \mathcal{L}] = \int d^3r [i\psi_\alpha^* \dot{\psi}_\alpha - i\bar{\psi}\gamma^\mu \partial_\mu \psi + m\bar{\psi}\psi] \\ &= \int d^3r \psi^\dagger \left[-i \sum_{k=1}^3 \gamma^0 \gamma^k \partial_k + m \right] \psi = \int d^3r \psi^\dagger \hat{\mathcal{H}} \psi \end{aligned} \quad (3.127)$$

où $\hat{\mathcal{H}}$ est l'opérateur différentiel introduit à l'éq. (3.93), à ne pas confondre avec la densité hamiltonienne \mathcal{H} (sans accent). Or, les solutions (3.105) et (3.107) obéissent précisément aux équations aux valeurs propres

$$\begin{aligned} \hat{\mathcal{H}} u_{\mathbf{p},s} e^{i\mathbf{p}\cdot\mathbf{r}} &= E_{\mathbf{p}} u_{\mathbf{p},s} e^{i\mathbf{p}\cdot\mathbf{r}} & (s=1,2) \\ \hat{\mathcal{H}} u_{\mathbf{p},s} e^{i\mathbf{p}\cdot\mathbf{r}} &= -E_{\mathbf{p}} u_{\mathbf{p},s} e^{i\mathbf{p}\cdot\mathbf{r}} & (s=3,4) \end{aligned} \quad (3.128)$$

En substituant le développement en modes (3.113), on trouve sans peine que

$$H = \sum_{\mathbf{p}} E_{\mathbf{p}} (c_{\mathbf{p},1}^\dagger c_{\mathbf{p},1} + c_{\mathbf{p},2}^\dagger c_{\mathbf{p},2} - c_{\mathbf{p},3}^\dagger c_{\mathbf{p},3} - c_{\mathbf{p},4}^\dagger c_{\mathbf{p},4}) \quad (3.129)$$

En d'autres termes, l'opérateur d'énergie est simplement la somme sur les modes des énergies de chaque mode, multipliées par le nombre d'électrons dans chaque mode.

De même, on peut définir un opérateur de la quantité de mouvement comme suit :

$$\mathbf{P} = \sum_{\mathbf{p},s} \mathbf{p} c_{\mathbf{p},s}^\dagger c_{\mathbf{p},s} \quad (3.130)$$

3.C.4 Antiparticules

Même après l'introduction des relations d'anticommutation, il existe un problème fondamental dans notre interprétation des états $c_{\mathbf{p},s}^\dagger|0\rangle$ en tant qu'électrons : l'existence des solutions à énergie négative, correspondant aux indices $s = 3, 4$. Il s'agit d'un problème en apparence très sérieux, car si on peut créer des états d'énergie négative en appliquant $c_{\mathbf{p},3}^\dagger$ et $c_{\mathbf{p},4}^\dagger$ sur le vide $|0\rangle$, c'est que le vide n'est pas l'état fondamental. Dans ce cas, l'état fondamental serait obtenu en occupant tous les états d'énergie négative, et il y en a deux pour chaque valeur de quantité de mouvement possible ! Cet état fondamental contiendrait un nombre infini de particules et serait l'analogue de ce qu'on appelle la mer de Fermi en physique de l'état solide. On l'appelle de fait la *mer de Dirac*, et on le note $|F\rangle$:

$$|F\rangle = \prod_{\mathbf{p}} \prod_{s=3,4} c_{\mathbf{p},s}^\dagger |0\rangle \quad (3.131)$$

Comment cet état peut-il être bien défini s'il y a un nombre infini de termes dans ce produit ? Fondamentalement, il ne peut y avoir un nombre infini de termes : l'infini mathématique est toujours une approximation en physique. On peut supposer que l'espace est fondamentalement discret et fini, donc que les vecteurs d'ondes sont en nombre fini.²

Annihiler une particule d'énergie négative de la mer de Dirac revient à augmenter l'énergie du système de $E_{\mathbf{p}}$ et modifier la quantité de mouvement du système de $-\mathbf{p}$. De plus, annihiler une particule de projection de spin $S_z = \frac{1}{2}$ revient à ajouter au total une projection de spin $S_z = -\frac{1}{2}$. On peut alors définir de nouveaux opérateurs ainsi :

$$d_{\mathbf{p},1} = c_{-\mathbf{p},4}^\dagger \quad \text{et} \quad d_{\mathbf{p},2} = c_{-\mathbf{p},3}^\dagger \quad (3.132)$$

Ces opérateurs obéissent aux mêmes relations d'anticommutation que les autres :

$$\{d_{\mathbf{p},s}, d_{\mathbf{p}',s'}^\dagger\} = \delta_{\mathbf{p},\mathbf{p}'} \delta_{ss'} \quad \{d_{\mathbf{p},s}, d_{\mathbf{p}',s'}\} = 0 \quad (3.133)$$

(ceci ne serait pas vrai pour des bosons, car un changement de signe des commutateurs serait survenu). L'état fondamental $|F\rangle$ respecte alors les conditions suivantes :

$$c_{\mathbf{p},1}|F\rangle = c_{\mathbf{p},2}|F\rangle = d_{\mathbf{p},1}|F\rangle = d_{\mathbf{p},2}|F\rangle = 0 \quad (3.134)$$

Le hamiltonien prend alors la forme suivante :

$$H = \sum_{\mathbf{p}} E_{\mathbf{p}} \left(c_{\mathbf{p},1}^\dagger c_{\mathbf{p},1} + c_{\mathbf{p},2}^\dagger c_{\mathbf{p},2} - d_{-\mathbf{p},1}^\dagger d_{-\mathbf{p},1} - d_{-\mathbf{p},2}^\dagger d_{-\mathbf{p},2} \right) \quad (3.135)$$

Comme $E_{\mathbf{p}} = E_{-\mathbf{p}}$, on peut changer l'indice de sommation \mathbf{p} en $-\mathbf{p}$ dans les deux derniers termes :

$$H = \sum_{\mathbf{p}} E_{\mathbf{p}} \left(c_{\mathbf{p},1}^\dagger c_{\mathbf{p},1} + c_{\mathbf{p},2}^\dagger c_{\mathbf{p},2} - d_{\mathbf{p},1}^\dagger d_{\mathbf{p},1} - d_{\mathbf{p},2}^\dagger d_{\mathbf{p},2} \right) \quad (3.136)$$

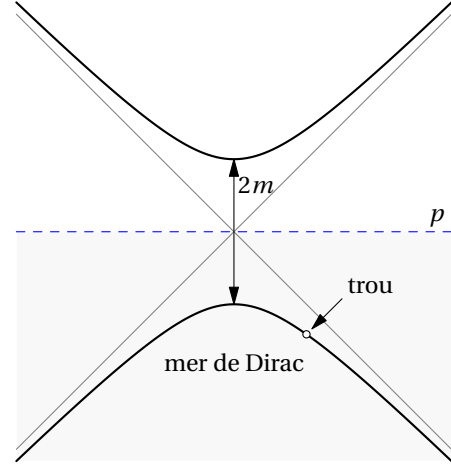
2. Ce besoin de supposer que l'ensemble des vecteurs d'onde possibles est fini se fera également sentir lorsque la théorie des perturbations sera appliquée à des particules en interaction, et constitue la première étape d'une procédure appelée *renormalisation*.

ce qui devient, en utilisant les relations d'anticommutation afin de changer l'ordre de $d_{\mathbf{p},s}$ et $d_{\mathbf{p},s}^\dagger$,

$$H = \sum_{\mathbf{p}} E_{\mathbf{p}} \left(c_{\mathbf{p},1}^\dagger c_{\mathbf{p},1} + c_{\mathbf{p},2}^\dagger c_{\mathbf{p},2} + d_{\mathbf{p},1}^\dagger d_{\mathbf{p},1} + d_{\mathbf{p},2}^\dagger d_{\mathbf{p},2} \right) - 2 \sum_{\mathbf{p}} E_{\mathbf{p}} \quad (3.137)$$

Le dernier terme est une constante, en fait l'énergie de l'état fondamental $|F\rangle$. On peut laisser tomber cette constante, car l'énergie est de toute manière définie à une constante près.³

FIGURE 3.1
Schéma de la mer de Dirac.



Comme les définitions (3.132) font intervenir un changement dans le signe de la quantité de mouvement, on préserve quand même la forme correcte de l'opérateur quantité de mouvement :

$$\begin{aligned} \mathbf{P} &= \sum_{\mathbf{p}} \mathbf{p} \left(c_{\mathbf{p},1}^\dagger c_{\mathbf{p},1} + c_{\mathbf{p},2}^\dagger c_{\mathbf{p},2} + d_{-\mathbf{p},1}^\dagger d_{-\mathbf{p},1} + d_{-\mathbf{p},2}^\dagger d_{-\mathbf{p},2} \right) \\ &= \sum_{\mathbf{p}} \mathbf{p} \left(c_{\mathbf{p},1}^\dagger c_{\mathbf{p},1} + c_{\mathbf{p},2}^\dagger c_{\mathbf{p},2} - d_{-\mathbf{p},1}^\dagger d_{-\mathbf{p},1} - d_{-\mathbf{p},2}^\dagger d_{-\mathbf{p},2} \right) \\ &= \sum_{\mathbf{p}} \mathbf{p} \left(c_{\mathbf{p},1}^\dagger c_{\mathbf{p},1} + c_{\mathbf{p},2}^\dagger c_{\mathbf{p},2} + d_{\mathbf{p},1}^\dagger d_{\mathbf{p},1} + d_{\mathbf{p},2}^\dagger d_{\mathbf{p},2} \right) \end{aligned} \quad (3.138)$$

À la deuxième ligne, nous avons utilisé les relations d'anticommutation; il n'y a pas de constante additive nette dans ce cas car les constantes associées à \mathbf{p} et $-\mathbf{p}$ se compensent exactement. À la troisième ligne, nous avons changé de variable de sommation $\mathbf{p} \rightarrow -\mathbf{p}$ pour les deux derniers termes, ce qui change leur signe en raison du préfacteur \mathbf{p} qui, lui aussi, change de signe.

Charge électrique Passons maintenant à la charge électrique. Le quadrivecteur courant associé à un spineur de Dirac est nécessairement de la forme

$$j^\mu = e \bar{\psi} \gamma^\mu \psi \quad (3.139)$$

car l'équation de continuité est alors une conséquence de l'équation de Dirac (3.83) :

$$\begin{aligned} \partial_\mu j^\mu &= e \partial_\mu \bar{\psi} \gamma^\mu \psi + e \bar{\psi} \gamma^\mu \partial_\mu \psi \\ &= i e m \bar{\psi} \psi - i e m \bar{\psi} \psi = 0 \end{aligned} \quad (3.140)$$

3. Sauf en relativité générale, où les équations d'Einstein font intervenir la valeur absolue de la densité d'énergie. Ce terme que nous négligeons allègrement ici contribue à ce qu'on appelle l'énergie sombre.

La charge électrique est alors

$$Q = \int d^3r j^0 = e \int d^3r \bar{\psi} \gamma^0 \psi = e \int d^3r \psi^\dagger \psi \quad (3.141)$$

En substituant le développement en modes (3.113) et en utilisant l'orthogonalité des différents spineurs, on trouve immédiatement l'expression suivante pour la charge électrique :

$$Q = e \sum_{\mathbf{p}} \left(c_{\mathbf{p},1}^\dagger c_{\mathbf{p},1} + c_{\mathbf{p},2}^\dagger c_{\mathbf{p},2} + c_{\mathbf{p},3}^\dagger c_{\mathbf{p},3} + c_{\mathbf{p},4}^\dagger c_{\mathbf{p},4} \right) \quad (3.142)$$

Après la transformation (3.132), la charge électrique s'exprime plutôt comme suit :

$$Q = e \sum_{\mathbf{p}} \left(c_{\mathbf{p},1}^\dagger c_{\mathbf{p},1} + c_{\mathbf{p},2}^\dagger c_{\mathbf{p},2} - d_{\mathbf{p},1}^\dagger d_{\mathbf{p},1} - d_{\mathbf{p},2}^\dagger d_{\mathbf{p},2} \right) + Q_0 \quad (3.143)$$

où Q_0 est la constante $\sum_{\mathbf{p}} 2$, soit la charge électrique contenue dans la mer de Dirac. Or l'univers est globalement neutre : la somme de toutes ces constantes Q_0 sur tous les types de fermions connus (leptons et quarks) doit être nulle.⁴

C'est ici que l'interprétation des solutions prend toute sa richesse : les particules créées par les opérateurs $d_{\mathbf{p},s}^\dagger$ ne sont pas des particules, mais des antiparticules ! Elles ont les mêmes propriétés que les particules créées par les opérateurs $c_{\mathbf{p},s}^\dagger$, mais une charge opposée. Autrement dit, les antiparticules contribuent de manière négative à la charge électrique.

La théorie de Dirac prédit donc l'existence d'antiparticules à l'électron, et en fait à toutes les particules de spin $\frac{1}{2}$. L'antiélectron porte le nom de *positron* et a été identifié dès 1932 par Carl Anderson.

Notons enfin que, en fonction des opérateurs de création et d'annihilation, le champ de Dirac a l'expression suivante :

$$\psi(\mathbf{r}) = \frac{1}{\sqrt{\mathcal{V}}} \sum_{\mathbf{p}} \sum_{s=1,2} \left(c_{\mathbf{p},s} u_{\mathbf{p},s} e^{i\mathbf{p}\cdot\mathbf{r}} + d_{\mathbf{p},s}^\dagger v_{\mathbf{p},s} e^{-i\mathbf{p}\cdot\mathbf{r}} \right) \quad (3.144)$$

où on a défini des spineurs appropriés aux antiparticules :

$$v_{\mathbf{p},1} = u_{-\mathbf{p},4} \quad v_{\mathbf{p},2} = u_{-\mathbf{p},3} \quad (3.145)$$

4. Si on consulte le tableau des particules élémentaires à la Fig. 1.1, on constate que ce n'est apparemment pas le cas, car la somme des charges électriques dans chaque famille de fermions n'est pas nulle. Mais ce n'est qu'une apparence, car chaque quark existe en trois variétés équivalentes, qu'on appelle *couleurs*, et cela restaure l'équilibre de la charge électrique dans chaque famille.

D Problèmes

Question 3.1 Matrices de Dirac

Démontrez les propriétés suivantes des matrices de Dirac, à l'aide de la représentation chirale (3.76) :

A

$$\gamma^\mu \gamma^\nu + \gamma^\nu \gamma^\mu = 2g^{\mu\nu} \quad (3.146)$$

B

$$\gamma^0 \gamma^\mu \gamma^0 = (\gamma^\mu)^\dagger \quad (3.147)$$

Problème 3.1 Tenseur antisymétrique

Montrez que si ψ est un spineur de Dirac, alors $\bar{\psi} \gamma^\mu \gamma^\nu \psi$ se comporte, lors d'une transformation de Lorentz, comme un tenseur de rang 2.

Problème 3.2 Démonstration différente de l'équation de Dirac

Essayons maintenant d'arriver à l'équation de Dirac d'une manière différente de celle que nous avons utilisée en classe. Dans l'espace réciproque (impulsion et énergie), l'équation de Klein-Gordon prend la forme

$$(g^{\mu\nu} p_\mu p_\nu - m^2) \psi = 0 \quad (3.148)$$

où $p_\mu = (E, -\mathbf{p})$ est le quadrivecteur d'énergie impulsion.

Dans le but d'avoir une équation du premier ordre en dérivées – et donc linéaire en p_μ – essayons de factoriser cette équation comme suit :

$$(g^{\mu\nu} p_\mu p_\nu - m^2) = (\beta^\mu p_\mu + m)(\gamma^\nu p_\nu - m) \quad (3.149)$$

où β^μ et γ^ν sont des quantités à déterminer.

A Démontrer que β^μ et γ^ν ne peuvent être des nombres ordinaires (réels ou complexes), c'est-à-dire que la factorisation ne peut pas fonctionner dans ce cas. Démontrez en outre que la factorisation fonctionne si les β^μ, γ^ν sont des matrices telles que $\beta^\mu = \gamma^\mu$ et

$$\gamma^\mu \gamma^\nu + \gamma^\nu \gamma^\mu = 2g^{\mu\nu} \quad (3.150)$$

B Démontrez qu'en dimension 3 d'espace-temps (donc deux dimensions d'espace), on pourrait choisir une représentation des matrices de Dirac de dimension 2.

C Démontrez qu'en quatre dimensions d'espace-temps, une représentation de dimension 2 est impossible.

Problème 3.3 Valeur moyenne du spin d'un électron

A Étant donné un spineur de Dirac ψ , comment peut-on calculer la valeur des composantes (s_x, s_y, s_z) du spin dans cet état?

B Considérons en particulier un spineur associé à un électron de quantité de mouvement $\mathbf{p} = p\mathbf{z}$, mais dont l'orientation de spin est quelconque, c'est-à-dire une superposition $\alpha u_1 + \beta u_2$, où $|\alpha|^2 + |\beta|^2 = 1$. Montrer que la projection du spin sur \mathbf{z} est $\frac{1}{2}(|\alpha|^2 - |\beta|^2)$, alors que les projections du spin le long de \mathbf{x} et \mathbf{y} sont

$$\frac{m}{E_p} \operatorname{Re}(\alpha^* \beta) \quad \text{et} \quad \frac{m}{E_p} \operatorname{Im}(\alpha^* \beta) \quad (3.151)$$

C Que conclure de ceci pour un électron ultrarelativiste?

Problème 3.4 Application de la covariance de l'équation de Dirac

À partir de la solution de l'équation de Dirac pour une particule libre au repos, trouvez une expression pour la solution correspondant à une particule d'impulsion \mathbf{p} en appliquant une transformation de Lorentz vers le référentiel approprié. Utilisez à cette fin les propriétés de transformation des spineurs droit et gauche sous transformation de Lorentz. N'oubliez pas cependant de travailler dans la représentation de Dirac et non la représentation chirale.

Problème 3.5 Paradoxe de Klein

Considérons une particule de masse m décrite par l'équation de Dirac, en présence d'une énergie potentielle constante V dans la région $z > 0$, et nulle dans la région $z < 0$. Cette énergie potentielle provient, par exemple, d'un potentiel électrique Φ constant dans la région $z > 0$ ($V = e\Phi$). La particule a un spin polarisé dans la direction z et se dirige dans la direction $+z$ avec une énergie totale E (incluant l'énergie de masse). À l'interface ($z = 0$), une onde réfléchie est générée, ainsi qu'une onde transmise. Il s'agit donc du problème de la réflexion et de la transmission par une marche de potentiel, en version relativiste.^a

A Comment doit-on adapter les solutions à l'équation de Dirac trouvées dans les notes pour les appliquer à la région $z > 0$?

B En vous servant des solutions d'ondes planes démontrées dans les notes, exprimez la forme

de ψ dans les deux régions. Vous pouvez supposer que l'amplitude de l'onde incidente est 1 (ou toute autre valeur); les amplitudes des deux autres ondes seront déterminées par la condition de continuité à $z = 0$. Supposez que l'interface à $z = 0$ n'affecte pas la projection de spin (est-ce justifié)?

C Calculez le coefficient de réflexion R (le rapport du courant réfléchi sur le courant incident) et le coefficient de transmission T (courant transmis sur courant incident). Montrez que $T + R = 1$.

D Montrez que (i) $R = 1$ si $|E - V| < m$, (ii) $R > 1$ si $V > E + m$, et (iii) $R < 1$ sinon. Faites (i) un graphique de R en fonction de E pour $V = m$ et $V = 3m$ (domaine de E : de m à $5m$) et (ii) un graphique de R en fonction de V pour $E = 3m/2$ et $E = 3m$. En quoi ces résultats sont-ils paradoxaux? Pouvez-vous expliquer la cause physique de ce paradoxe?

a. O. Klein, *Die reflexion von Elektronen an einem Potentialsprung nach der relativischen Dynamik von Dirac*, Zeitschrift für Physik **53**, 157-165 (1929).

CHAPITRE 4

L'ÉLECTRODYNAMIQUE QUANTIQUE

A Le champ électromagnétique

4.A.1 Particule chargée dans un champ électromagnétique

Rappelons les grandes lignes de la description classique de l'interaction entre une particule de charge e et le champ électromagnétique. On sait qu'à l'action S_0 d'une particule libre on doit ajouter une action S_1 décrivant l'interaction de cette particule avec le champ électromagnétique, et que cette action prend la forme suivante :

$$S_1 = -e \int A_\mu dx^\mu \quad (4.1)$$

où A_μ est le quadrivecteur du potentiel électromagnétique et l'intégrale est effectuée sur la trajectoire de la particule. Comme l'intégrand est la contraction de deux quadrivecteurs, il est forcément invariant de Lorentz. L'action (4.1) a aussi la propriété d'*invariance de jauge*. Si on modifie le potentiel électromagnétique en lui ajoutant le gradient d'une fonction quelconque ξ de l'espace-temps :

$$A_\mu \rightarrow A'_\mu = A_\mu + \partial_\mu \xi \quad (4.2)$$

alors la variation de l'action ne change pas. En effet, le changement dans l'action qui résulte d'une transformation de jauge est

$$S'_1 = S_1 - e \int \partial_\mu \xi dx^\mu = S_1 - e \xi(x_f) + e \xi(x_i) \quad (4.3)$$

où x_i et x_f représente les extrémités du mouvement. Or, lorsque le principe de moindre action est appliqué, les extrémités du mouvement sont supposées fixes. Donc les variations δS_1 et $\delta S'_1$ sont identiques, ainsi que les équations du mouvement qui en découlent.

En fonction des composantes temporelles et spatiales (Φ, \mathbf{A}) du potentiel électromagnétique, le lagrangien d'une particule devient

$$L = -m \sqrt{1 - v^2} + e \mathbf{A} \cdot \mathbf{v} - e \Phi \quad (4.4)$$

Le moment conjugué à la position \mathbf{r} est alors

$$\mathbf{p} = \frac{\partial L}{\partial \mathbf{v}} = m\mathbf{v} + e\mathbf{A} \quad (4.5)$$

et le hamiltonien

$$H = \mathbf{v} \cdot \mathbf{p} - L = \frac{m}{\sqrt{1-v^2}} + e\Phi \quad (4.6)$$

où il est compris que la vitesse \mathbf{v} doit être remplacée par $(\mathbf{p} - e\mathbf{A})/m\gamma$, ce qui donne

$$H = \sqrt{(\mathbf{p} - e\mathbf{A})^2 + m^2} + e\Phi \quad (4.7)$$

On montre facilement que l'action (4.1) mène directement à l'équation du mouvement

$$\frac{d\mathbf{p}}{dt} = e\mathbf{E} + e\mathbf{v} \wedge \mathbf{B} \quad (4.8)$$

où les champs électrique et magnétique sont reliés de la manière suivante aux potentiels électromagnétiques :

$$\mathbf{E} = -\nabla\Phi - \frac{\partial \mathbf{A}}{\partial t} \quad \mathbf{B} = \nabla \wedge \mathbf{A} \quad (4.9)$$

Une expression covariante de la même équation du mouvement est

$$m \ddot{x}_\mu = e F_{\mu\nu} \dot{x}^\nu \quad (4.10)$$

où on a introduit le *tenseur de Faraday*

$$F_{\mu\nu} = \partial_\mu A_\nu - \partial_\nu A_\mu \quad (4.11)$$

dont les composantes explicites sont :

$$F_{\mu\nu} : \begin{pmatrix} 0 & E_x & E_y & E_z \\ -E_x & 0 & -B_z & B_y \\ -E_y & B_z & 0 & -B_x \\ -E_z & -B_y & B_x & 0 \end{pmatrix} \quad (4.12)$$

Notons que le tenseur de Faraday est invariant de jauge, comme on le voit facilement en appliquant la transformation (4.2). C'est d'ailleurs l'expression locale et invariante de jauge la plus simple impliquant le potentiel électromagnétique.

4.A.2 Action du champ électromagnétique et équations de Maxwell

Les équations de Maxwell, qui gouvernent le champ électromagnétique, peuvent aussi être dérivées du principe de la moindre action. L'action du champ électromagnétique doit s'exprimer en fonction du potentiel A_μ et de ses dérivées :

$$S = \int d^4x \mathcal{L}(A_\mu, \partial_\nu A_\mu) \quad (4.13)$$

Cette action doit contenir deux termes : l'un décrivant la dynamique propre du champ et l'autre décrivant l'interaction du champ avec la matière chargée. Ce dernier terme nous est déjà connu : il s'agit de l'action (4.1) ci-haut. En exprimant celui-ci en fonction du quadrivecteur de densité de courant au lieu de la quadrivitesse, on trouve

$$S_1 = - \int d^4x J^\mu A_\mu \quad (4.14)$$

En effet, le quadrivecteur courant peut s'exprimer comme à l'éq. (1.29) en fonction de particules ponctuelles. On trouve alors

$$S_1 = - \int d^4x J^\mu A_\mu = - \sum_i q_i \int d^4x A_\mu u^{(i)\mu} \delta(\mathbf{r} - \mathbf{r}^{(i)}) / \gamma^{(i)} \quad (4.15)$$

mais l'intégrale sur les coordonnées spatiales peut se faire à l'aide des fonctions delta, laissant

$$S_1 = - \sum_i q_i \int \frac{dt}{\gamma^{(i)}} u^{(i)\mu} A_\mu = - \sum_i q_i \int d\tau u^{(i)\mu} A_\mu = - \sum_i q_i \int dx^{(i)\mu} A_\mu \quad (4.16)$$

Ce qui correspond à l'action (4.1) pour un ensemble de particules ponctuelles.

Il reste à trouver S_0 , l'action propre du champ, ou la densité lagrangienne correspondante \mathcal{L}_0 . Sachant que les équations de Maxwell sont linéaires, il faut que \mathcal{L}_0 soit au plus quadratique en fonction des champs. Il faut également que \mathcal{L}_0 soit (i) invariant de Lorentz et (ii) invariant de jauge. La seule possibilité est la suivante :

$$\mathcal{L}_0 = -\frac{1}{4} F^{\mu\nu} F_{\mu\nu} = -\frac{1}{2} F^{\mu\nu} \partial_\mu A_\nu \quad (4.17)$$

où $F_{\mu\nu}$ est le tenseur de Faraday. Le facteur de $-1/4$ a été ajouté dans le but de retrouver les équations de Maxwell dans leur normalisation habituelle par rapport au courant. La densité lagrangienne totale est donc

$$\mathcal{L} = -J^\mu A_\mu - \frac{1}{2} F^{\mu\nu} \partial_\mu A_\nu \quad (4.18)$$

Utilisons maintenant le principe de la moindre action pour obtenir les équations du mouvement. Il faut pour cela ajouter une variation δA_μ au quadrivecteur du potentiel et s'assurer que la variation correspondante de l'action est nulle au premier ordre. Notons que

$$\delta(F^{\mu\nu} F_{\mu\nu}) = 2F^{\mu\nu} \delta F_{\mu\nu} = 4F^{\mu\nu} \partial_\mu \delta A_\nu \quad (4.19)$$

Donc la variation de l'action est

$$\delta S = \int d^4x \{ -J^\nu \delta A_\nu - F^{\mu\nu} \partial_\mu \delta A_\nu \} \quad (4.20)$$

Le deuxième terme peut être intégré par parties, de la manière suivante : considérant que

$$F^{\mu\nu} \partial_\mu \delta A_\nu = \partial_\mu (F^{\mu\nu} \delta A_\nu) - \partial_\mu F^{\mu\nu} \delta A_\nu, \quad (4.21)$$

le premier terme est une dérivée totale (dans ce cas-ci, une quadridivergence), qui ne contribue pas à la variation de l'action. Il reste donc

$$\delta S = \int d^4x \{ -J^\nu + \partial_\mu F^{\mu\nu} \} \delta A_\nu \quad (4.22)$$

Si cette variation s'annule pour n'importe quelle variation δA_ν , il faut que le tenseur électromagnétique satisfasse à l'équation suivante :

$$\partial_\mu F^{\mu\nu} = J^\nu \quad (4.23)$$

Cette équation constitue une partie des équations de Maxwell : la loi de Gauss et la loi d'Ampère. En fonction des composantes électriques et magnétiques du tenseur de Faraday, elle s'exprime ainsi :

$$\nabla \cdot \mathbf{E} = \rho \quad \nabla \wedge \mathbf{B} - \frac{\partial \mathbf{E}}{\partial t} = \mathbf{J} \quad (4.24)$$

L'autre partie provient de la définition même de $F_{\mu\nu}$ en fonction de A_μ :

$$\varepsilon_{\mu\nu\lambda\rho} \partial^\nu F^{\lambda\rho} = 0 \quad (4.25)$$

ce qui s'exprime également comme suit :

$$\nabla \cdot \mathbf{B} = 0 \quad \nabla \wedge \mathbf{E} + \frac{\partial \mathbf{B}}{\partial t} = 0 \quad (4.26)$$

On reconnaît la loi qui représente l'absence de charge magnétique ainsi que la loi de Faraday.

4.A.3 Invariance de jauge en mécanique quantique

Le hamiltonien (4.7) d'une particule de charge e en présence d'un champ électromagnétique s'obtient de celui d'une particule libre en suivant la prescription suivante :

$$\mathbf{p} \rightarrow \mathbf{p} - e\mathbf{A} \quad \text{et} \quad H \rightarrow H - e\Phi \quad (4.27)$$

ou encore, en langage covariant,

$$p_\mu \rightarrow p_\mu - eA_\mu \quad (4.28)$$

De par la correspondance $p_\mu \rightarrow i\partial_\mu$, cela revient à faire les substitutions suivantes dans l'équation de Schrödinger ou l'équation de Dirac :

$$\partial_\mu \rightarrow \mathcal{D}_\mu := \partial_\mu + ieA_\mu \quad \text{ou} \quad \begin{cases} \nabla \rightarrow \mathcal{D} := \nabla - ie\mathbf{A} \\ \frac{\partial}{\partial t} \rightarrow \mathcal{D}_t := \frac{\partial}{\partial t} + ie\Phi \end{cases} \quad (4.29)$$

L'opérateur $\mathcal{D}_\mu = (\mathcal{D}_t, \mathcal{D})$ est appelée *dérivée covariante*. En fonction de ces opérateurs différentiels, l'équation de Schrödinger prend la forme suivante :

$$i\mathcal{D}_t \psi = -\frac{1}{2m} \mathcal{D}^2 \psi \quad (4.30)$$

et l'équation de Dirac la forme suivante :

$$i\gamma^\mu \mathcal{D}_\mu \psi - m\psi = 0 \quad (4.31)$$

La prescription qui consiste à remplacer les dérivées ordinaires par des dérivées covariantes porte le nom de *couplage minimal*. Elle garantit l'invariance des équations d'onde par rapport aux transformations de jauge.

Afin qu'une équation différentielle pour une quantité ψ (telle la fonction d'onde) soit invariante de jauge, il suffit que le couplage minimal soit utilisé et que la fonction ψ se transforme ainsi lors d'une transformation de jauge :

$$\psi \rightarrow e^{-ie\xi} \psi \quad (4.32)$$

Comme ξ dépend de la position et du temps en général, il s'agit d'un changement de phase *local*, par opposition à *global*. La propriété principale des dérivées covariantes est la *covariance* sous transformation de jauge, c'est-à-dire que, lors d'une transformation de jauge, les dérivées covariantes sont modifiées par un simple facteur de phase local :

$$\mathcal{D}_\mu \psi \rightarrow e^{-ie\xi} \mathcal{D}_\mu \psi \quad (4.33)$$

Il est alors manifeste que les équations de Schrödinger et de Dirac sont invariantes de jauge.

La relation (4.33) se démontre facilement :

$$\begin{aligned} (\mathcal{D}_\mu \psi)' &= (\partial_\mu + ieA'_\mu)(e^{-ie\xi} \psi) \\ &= e^{-ie\xi} \partial_\mu \psi - ie \partial_\mu \xi e^{-ie\xi} \psi + ieA'_\mu e^{-ie\xi} \psi \\ &= e^{-ie\xi} \partial_\mu \psi + ieA_\mu e^{-ie\xi} \psi \\ &= e^{-ie\xi} (\partial_\mu + ieA_\mu) \psi = e^{-ie\xi} \mathcal{D}_\mu \psi \end{aligned} \quad (4.34)$$

où nous avons utilisé la relation $A'_\mu = A_\mu + \partial_\mu \xi$ à la troisième ligne.

On sait qu'en mécanique quantique la fonction d'onde peut être multipliée par une phase *globale* sans en modifier le sens physique. On apprend ici que la fonction d'onde d'une particule chargée peut être multipliée par une phase *locale* sans non plus modifier son sens physique, pourvu que cette multiplication s'accompagne d'un changement des potentiels électromagnétiques, tel qu'en l'éq. (4.2). Le fait de pouvoir transformer une invariance globale en invariance locale grâce à la présence de potentiels scalaire et vecteur est à la base des *théories de jauge* qui décrivent les interactions forte (QCD), faible et électromagnétique.

4.A.4 Couplage minimal

Appliquons le couplage minimal à l'équation de Dirac $(i\gamma^\mu \partial_\mu - m)\psi = 0$, c'est-à-dire remplaçons la dérivée ordinaire ∂_μ par une dérivée covariante \mathcal{D}_μ . Rappelons que cette prescription garantit l'invariance de jauge de la théorie considérée. On trouve alors

$$(i\gamma^\mu \mathcal{D}_\mu - m)\psi = (i\gamma^\mu \partial_\mu - eA_\mu \gamma^\mu - m)\psi = 0 \quad (4.35)$$

Le lagrangien associé à cette équation est alors

$$\mathcal{L}_D = \bar{\psi}(i\gamma^\mu \mathcal{D}_\mu - m)\psi = \bar{\psi}(i\gamma^\mu \partial_\mu - eA_\mu \gamma^\mu - m)\psi \quad (4.36)$$

Le couplage du champ de Dirac avec le champ électromagnétique prend la forme

$$\mathcal{L}_1 = -A_\mu J^\mu \quad J^\mu = e \bar{\psi} \gamma^\mu \psi \quad (4.37)$$

où J^μ est le quadrivecteur courant associé aux particules décrites par ψ .

Solution de l'équation de Dirac pour l'atome d'hydrogène

Il est possible de résoudre l'équation de Dirac en présence d'un champ électromagnétique (4.35) dans quelques cas, notamment en présence d'un champ coulombien, comme dans l'atome d'hydrogène¹. On trouve dans ce cas les niveaux d'énergie suivants :

$$E_{n,j} = m \left(1 + \frac{\alpha^2}{\left(n - j - \frac{1}{2} + \sqrt{(j + 1/2)^2 - \alpha^2} \right)^2} \right)^{-1/2} \quad (4.38)$$

où n est le nombre quantique principal (un entier positif), alors que les valeurs possibles de j sont $\frac{1}{2}, \frac{3}{2}, \dots$. En fait, $j = l \pm \frac{1}{2}$ où $l \in \{0, 1, \dots, n-1\}$ est le nombre quantique orbital (la possibilité $j = l - \frac{1}{2}$ n'existe que pour $l > 0$). Ce résultat est en excellent accord avec la structure fine du spectre de l'atome d'hydrogène. En développant cette expression à l'ordre α^4 , on trouve

$$E_{n,j} = m - \frac{m\alpha^2}{2n^2} - \frac{m\alpha^4}{2n^3} \left(\frac{1}{j + 1/2} - \frac{3}{4n} \right) \quad (4.39)$$

Le premier terme est l'énergie de masse de l'électron. Le deuxième décrit les niveaux d'énergie de l'atome de Bohr, et le troisième la structure fine.

La première déviation (petite) expérimentalement observable entre la formule (4.38) et l'expérience sera le déplacement de Lamb, observé en 1946, et attribuable aux corrections radiatives (voir plus bas).

B Photons

4.B.1 Quantification du champ électromagnétique

Jauge de Coulomb La procédure de quantification du champ scalaire, suivie à la section 2.B, peut être appliquée au champ électromagnétique, à la différence que l'invariance de jauge pose une difficulté particulière. Il est préférable dans ce cas de *fixer une jauge*, c'est-à-dire d'imposer une condition particulière au potentiel A_μ de manière à simplifier la quantification. Nous allons imposer la condition suivante :

$$\nabla \cdot \mathbf{A} = 0 \quad (4.40)$$

1. C.G. Darwin, *Proc. Roy. Soc.* **A118**, 654 (1928) et W. Gordon, *Z. f. Phys.* **48**, 11 (1928).

qu'on appelle la *jauge de Coulomb* ou encore *jauge transverse*. Si le potentiel vecteur ne respecte pas cette condition d'emblée, alors il faut appliquer une transformation de jauge (4.2) telle que le nouveau potentiel \mathbf{A}' la respecte. Il faut donc trouver une fonction ξ telle que

$$\nabla \cdot \mathbf{A}' = \nabla \cdot \mathbf{A} - \nabla^2 \xi = 0 \implies \nabla^2 \xi = \nabla \cdot \mathbf{A} \quad (4.41)$$

\mathbf{A} étant une quantité connue, cette équation possède une solution explicite :

$$\xi(\mathbf{r}, t) = -\frac{1}{4\pi} \int d^3 r' \frac{\nabla \cdot \mathbf{A}(\mathbf{r}', t)}{|\mathbf{r} - \mathbf{r}'|} \quad (4.42)$$

Il est donc toujours possible d'imposer la condition (4.40).

Cette condition étant imposée, la loi de Gauss s'exprime comme suit en l'absence de sources :

$$\nabla \cdot \mathbf{E} = 0 \implies \nabla^2 \Phi + \frac{\partial}{\partial t} \nabla \cdot \mathbf{A} = 0 \implies \nabla^2 \Phi = 0 \quad (4.43)$$

La solution à cette équation dans l'espace infini est simplement $\Phi = 0$.

Lagrangien Le lagrangien du champ électromagnétique libre, c'est-à-dire en l'absence de sources, s'écrit explicitement ainsi :

$$\mathcal{L} = -\frac{1}{4} F^{\mu\nu} F_{\mu\nu} = \frac{1}{2} (\mathbf{E}^2 - \mathbf{B}^2) = \frac{1}{2} \left[\left(\frac{\partial \mathbf{A}}{\partial t} \right)^2 - (\nabla \wedge \mathbf{A})^2 \right] \quad (4.44)$$

Si on suit la procédure habituelle pour passer du lagrangien au hamiltonien, on trouve le moment conjugué

$$\boldsymbol{\pi}(\mathbf{r}) = \frac{\partial \mathcal{L}}{\partial \dot{\mathbf{A}}(\mathbf{r})} = \frac{\partial \mathbf{A}}{\partial t} \quad (4.45)$$

et la densité hamiltonienne

$$\begin{aligned} \mathcal{H} &= \boldsymbol{\pi}(\mathbf{r}) \cdot \frac{\partial \mathbf{A}}{\partial t} - \mathcal{L} \\ &= \frac{1}{2} \left[\left(\frac{\partial \mathbf{A}}{\partial t} \right)^2 + (\nabla \wedge \mathbf{A})^2 \right] \\ &= \frac{1}{2} [\boldsymbol{\pi}^2 + (\nabla \wedge \mathbf{A})^2] \\ &= \frac{1}{2} [\mathbf{E}^2 + \mathbf{B}^2] \end{aligned} \quad (4.46)$$

Transformées de Fourier Introduisons maintenant les transformées de Fourier :

$$\mathbf{A}(\mathbf{r}) = \frac{1}{\sqrt{\mathcal{V}}} \sum_{\mathbf{q}} \mathbf{A}_{\mathbf{q}} e^{i\mathbf{q} \cdot \mathbf{r}} \quad \mathbf{A}_{\mathbf{q}} = \frac{1}{\sqrt{\mathcal{V}}} \int d^3 r \mathbf{A}(\mathbf{r}) e^{-i\mathbf{q} \cdot \mathbf{r}} \quad (4.47)$$

et de même pour toute autre quantité (\mathbf{E} , \mathbf{B} , etc.). Il est sous-entendu ici que $\mathbf{A}(\mathbf{r})$ et $\mathbf{A}_{\mathbf{q}}$ dépendent du temps. Comme le potentiel vecteur est réel (par opposition à complexe), on voit aisément que

$$\mathbf{A}_{\mathbf{q}}^* = \mathbf{A}_{-\mathbf{q}} \quad (4.48)$$

et il en est de même de la transformée de Fourier de tout champ réel. La condition (4.40) s'exprime simplement en fonction des transformées de Fourier :

$$\mathbf{q} \cdot \mathbf{A}_{\mathbf{q}} = 0 \quad (4.49)$$

Ceci entraîne que, pour chaque vecteur d'onde \mathbf{q} , le vecteur $\mathbf{A}_{\mathbf{q}}$ est perpendiculaire à \mathbf{q} et donc n'a que deux composantes non nulles. C'est pour cette raison que la jauge de Coulomb est également appelée «jauge transverse».

En fonction des composantes de Fourier, le lagrangien s'écrit ainsi :

$$L = \frac{1}{2} \sum_{\mathbf{q}} \left[\frac{\partial \mathbf{A}_{\mathbf{q}}^*}{\partial t} \cdot \frac{\partial \mathbf{A}_{\mathbf{q}}}{\partial t} - (\mathbf{q} \wedge \mathbf{A}_{\mathbf{q}})^* \cdot (\mathbf{q} \wedge \mathbf{A}_{\mathbf{q}}) \right] \quad (4.50)$$

Remarquons cependant que

$$(\mathbf{q} \wedge \mathbf{A}_{\mathbf{q}})^* \cdot (\mathbf{q} \wedge \mathbf{A}_{\mathbf{q}}) = \mathbf{q}^2 \mathbf{A}_{\mathbf{q}}^* \cdot \mathbf{A}_{\mathbf{q}} - (\mathbf{q} \cdot \mathbf{A})^* (\mathbf{q} \cdot \mathbf{A}) \quad (4.51)$$

La condition transverse (4.49) nous permet donc d'écrire le lagrangien suivant :

$$L = \frac{1}{2} \sum_{\mathbf{q}} \left[\frac{\partial \mathbf{A}_{\mathbf{q}}^*}{\partial t} \cdot \frac{\partial \mathbf{A}_{\mathbf{q}}}{\partial t} - \mathbf{q}^2 \mathbf{A}_{\mathbf{q}}^* \cdot \mathbf{A}_{\mathbf{q}} \right] \quad (4.52)$$

Nous pouvons à ce stade introduire, pour chaque vecteur d'onde, une base de trois vecteurs ortho-normés $\mathbf{e}_{j\mathbf{q}}$ ($j = 1, 2, 3$) telle que $\mathbf{e}_{3\mathbf{q}} = \hat{\mathbf{q}}$, le vecteur unitaire dans la direction \mathbf{q} . Le potentiel vecteur $\mathbf{A}_{\mathbf{q}}$ admet donc un développement sur les deux premiers vecteurs de la base :

$$\mathbf{A}_{\mathbf{q}} = A_{1\mathbf{q}} \mathbf{e}_{1\mathbf{q}} + A_{2\mathbf{q}} \mathbf{e}_{2\mathbf{q}} \quad (4.53)$$

En fonction de ces deux composantes transverses, le lagrangien s'écrit enfin

$$L = \frac{1}{2} \sum_{\mathbf{q}} \sum_{j=1,2} \left[\frac{\partial A_{j\mathbf{q}}^*}{\partial t} \frac{\partial A_{j\mathbf{q}}}{\partial t} - \omega_q^2 A_{j\mathbf{q}}^* A_{j\mathbf{q}} \right] \quad \omega_q := |\mathbf{q}| \quad (4.54)$$

Sous cette forme, le lagrangien du champ électromagnétique est très semblable à celui d'un champ scalaire (voir éq. (2.57)). Si on se reporte à la section 2.B, on constate qu'on a ici l'équivalent de deux champs scalaires A_1 et A_2 , et qu'on peut donc introduire deux types d'opérateurs d'échelle indépendants $a_{1\mathbf{q}}$ et $a_{2\mathbf{q}}$, avec les relations de commutation

$$[a_{j\mathbf{q}}, a_{j'\mathbf{q}'}^\dagger] = \delta_{\mathbf{q}\mathbf{q}'} \delta_{jj'} \quad (4.55)$$

Le hamiltonien s'exprime ainsi en fonction de ces opérateurs :

$$H = \sum_{\mathbf{q}, j} \omega_q \left(a_{j\mathbf{q}}^\dagger a_{j\mathbf{q}} + \frac{1}{2} \right) \quad (4.56)$$

La composante de Fourier du potentiel vecteur s'exprime ainsi :

$$\mathbf{A}_{\mathbf{q}} = \frac{1}{\sqrt{2\omega_q}} \sum_{j=1,2} (a_{j\mathbf{q}} \mathbf{e}_{j\mathbf{q}} + a_{j-\mathbf{q}}^\dagger \mathbf{e}_{j\mathbf{q}}^*) \quad (4.57)$$

ce qui mène à l'expression suivante du potentiel vecteur :

$$\mathbf{A}(\mathbf{r}) = \sum_{\mathbf{q}, j} \frac{1}{\sqrt{2\omega_q \mathcal{V}}} \left(a_{j\mathbf{q}} \mathbf{e}_{j\mathbf{q}} e^{i\mathbf{q}\cdot\mathbf{r}} + a_{j-\mathbf{q}}^\dagger \mathbf{e}_{j\mathbf{q}}^* e^{-i\mathbf{q}\cdot\mathbf{r}} \right) \quad (4.58)$$

4.B.2 Photons

L'interprétation de ce résultat est identique à celle qui a été établie pour les quantas du champ scalaire. À chaque vecteur d'onde \mathbf{q} et polarisation $j = 1, 2$ du champ électromagnétique correspond un oscillateur harmonique de fréquence $\omega_q = q = |\mathbf{q}|$. L'état fondamental du champ électromagnétique est le produit tensoriel des états fondamentaux de tous les oscillateurs. On le note encore $|0\rangle$. L'opérateur

$$N_{j\mathbf{q}} = a_{j\mathbf{q}}^\dagger a_{j\mathbf{q}} \quad (4.59)$$

représente le nombre de photons de vecteur d'onde \mathbf{q} et de polarisation j . L'énergie du champ est alors la somme des énergies des photons :

$$H = \sum_{j,\mathbf{q}} \omega_q N_{j\mathbf{q}} \quad (4.60)$$

Il est également simple de démontrer que la quantité de mouvement du champ, définie comme

$$\mathbf{P} = \int d^3r \mathbf{E} \wedge \mathbf{B} \quad (4.61)$$

s'exprime ainsi en fonction des opérateurs de nombre :

$$\mathbf{P} = \sum_{j,\mathbf{q}} \mathbf{q} N_{j\mathbf{q}} \quad (4.62)$$

Donc les quantités physiques (énergie, quantité de mouvement) sont quantifiés, et $N_{j\mathbf{q}}$ représente le nombre de ces quantas associés à un vecteur d'onde et une polarisation donnés, donc le nombre de photons. La différence essentielle entre le champ électromagnétique et le champ scalaire est la nature vectorielle du premier, qui entraîne l'existence de polarisations. Par contre, la condition (4.49) interdit la polarisation longitudinale (dans la direction du vecteur d'onde) et il ne reste que deux polarisations transverses.

C L'électrodynamique quantique

L'électrodynamique quantique est la théorie quantique du champ électromagnétique et des particules chargées, en particulier décrites par l'équation de Dirac.

4.C.1 Hamiltonien d'interaction électron-photon

Le hamiltonien d'interaction électron-photon est l'opposé du lagrangien d'interaction, soit

$$H = e \int d^3 r A_\mu \bar{\psi} \gamma^\mu \psi \quad (4.63)$$

Dans la jauge de Coulomb, $A_0 = 0$ et donc

$$H = -e \int d^3 r \mathbf{A} \cdot \bar{\psi} \boldsymbol{\gamma} \psi \quad (4.64)$$

En substituant les développements en mode (3.144) et (4.58), on trouve

$$H = -\frac{e}{\mathcal{V}} \int d^3 r \sum_{j,\mathbf{k}} \sum_{s,s',\mathbf{p},\mathbf{p}'} \frac{1}{\sqrt{2\omega_k \mathcal{V}}} \left(a_{j\mathbf{k}} \boldsymbol{\epsilon}_{j\mathbf{k}} e^{i\mathbf{k}\cdot\mathbf{r}} + a_{j\mathbf{k}}^\dagger \boldsymbol{\epsilon}_{j\mathbf{k}}^* e^{-i\mathbf{k}\cdot\mathbf{r}} \right) \cdot \left(c_{\mathbf{p},s}^\dagger \bar{u}_{\mathbf{p},s} e^{-i\mathbf{p}\cdot\mathbf{r}} + d_{\mathbf{p},s} \bar{v}_{\mathbf{p},s} e^{i\mathbf{p}\cdot\mathbf{r}} \right) \boldsymbol{\gamma} \left(c_{\mathbf{p}',s'} u_{\mathbf{p}',s'} e^{i\mathbf{p}'\cdot\mathbf{r}} + d_{\mathbf{p}',s'}^\dagger v_{\mathbf{p}',s'} e^{-i\mathbf{p}'\cdot\mathbf{r}} \right) \quad (4.65)$$

Le développement s'effectue comme suit :

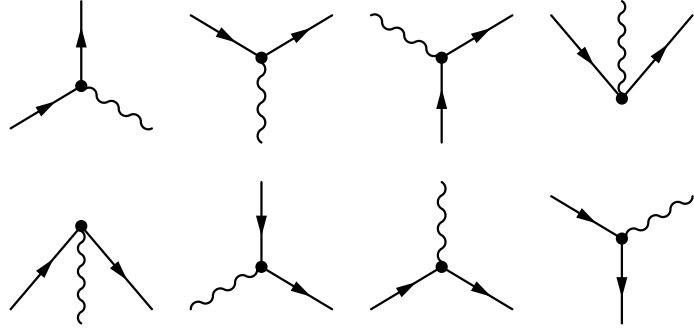
$$H_I = -e \int d^3 r \sum_{\mathbf{p},\mathbf{p}',\mathbf{q}} \sum_{s,s',r} \frac{1}{\mathcal{V}} \frac{1}{\sqrt{2\omega_q \mathcal{V}}} \left\{ \begin{aligned} & + c_{\mathbf{p},s}^\dagger c_{\mathbf{p}',s'} a_{r\mathbf{q}} (\bar{u}_{\mathbf{p},s} \boldsymbol{\gamma} u_{\mathbf{p}',s'}) \cdot \boldsymbol{\epsilon}_{r,\mathbf{q}} e^{i\mathbf{r}\cdot(-\mathbf{p}+\mathbf{p}'+\mathbf{q})} + c_{\mathbf{p},s}^\dagger d_{\mathbf{p}',s'}^\dagger a_{r\mathbf{q}} (\bar{u}_{\mathbf{p},s} \boldsymbol{\gamma} v_{\mathbf{p}',s'}) \cdot \boldsymbol{\epsilon}_{r,\mathbf{q}} e^{i\mathbf{r}\cdot(-\mathbf{p}-\mathbf{p}'+\mathbf{q})} \\ & + c_{\mathbf{p},s}^\dagger c_{\mathbf{p}',s'} a_{r\mathbf{q}}^\dagger (\bar{u}_{\mathbf{p},s} \boldsymbol{\gamma} u_{\mathbf{p}',s'}) \cdot \boldsymbol{\epsilon}_{r,\mathbf{q}}^* e^{i\mathbf{r}\cdot(-\mathbf{p}+\mathbf{p}'-\mathbf{q})} + c_{\mathbf{p},s}^\dagger d_{\mathbf{p}',s'}^\dagger a_{r\mathbf{q}}^\dagger (\bar{u}_{\mathbf{p},s} \boldsymbol{\gamma} v_{\mathbf{p}',s'}) \cdot \boldsymbol{\epsilon}_{r,\mathbf{q}}^* e^{i\mathbf{r}\cdot(-\mathbf{p}-\mathbf{p}'-\mathbf{q})} \\ & + d_{\mathbf{p},s} c_{\mathbf{p}',s'} a_{r\mathbf{q}} (\bar{v}_{\mathbf{p},s} \boldsymbol{\gamma} u_{\mathbf{p}',s'}) \cdot \boldsymbol{\epsilon}_{r,\mathbf{q}} e^{i\mathbf{r}\cdot(\mathbf{p}+\mathbf{p}'+\mathbf{q})} + d_{\mathbf{p},s} d_{\mathbf{p}',s'}^\dagger a_{r\mathbf{q}} (\bar{v}_{\mathbf{p},s} \boldsymbol{\gamma} v_{\mathbf{p}',s'}) \cdot \boldsymbol{\epsilon}_{r,\mathbf{q}} e^{i\mathbf{r}\cdot(\mathbf{p}-\mathbf{p}'+\mathbf{q})} \\ & + d_{\mathbf{p},s} c_{\mathbf{p}',s'} a_{r\mathbf{q}}^\dagger (\bar{v}_{\mathbf{p},s} \boldsymbol{\gamma} u_{\mathbf{p}',s'}) \cdot \boldsymbol{\epsilon}_{r,\mathbf{q}}^* e^{i\mathbf{r}\cdot(\mathbf{p}+\mathbf{p}'-\mathbf{q})} + d_{\mathbf{p},s} d_{\mathbf{p}',s'}^\dagger a_{r\mathbf{q}}^\dagger (\bar{v}_{\mathbf{p},s} \boldsymbol{\gamma} v_{\mathbf{p}',s'}) \cdot \boldsymbol{\epsilon}_{r,\mathbf{q}}^* e^{i\mathbf{r}\cdot(\mathbf{p}-\mathbf{p}'-\mathbf{q})} \end{aligned} \right\} \quad (4.66)$$

L'intégrale sur \mathbf{r} permet d'éliminer l'un des vecteurs d'ondes :

$$H_I = -e \sum_{\mathbf{p},\mathbf{q}} \sum_{s,s',r} \frac{1}{\sqrt{2\omega_q \mathcal{V}}} \left\{ \begin{aligned} & + c_{\mathbf{p},s}^\dagger c_{\mathbf{p}-\mathbf{q},s'} a_{r\mathbf{q}} (\bar{u}_{\mathbf{p},s} \boldsymbol{\gamma} u_{\mathbf{p}-\mathbf{q},s'}) \cdot \boldsymbol{\epsilon}_{r,\mathbf{q}} + c_{\mathbf{p},s}^\dagger d_{-\mathbf{p}+\mathbf{q},s'}^\dagger a_{r\mathbf{q}} (\bar{u}_{\mathbf{p},s} \boldsymbol{\gamma} v_{-\mathbf{p}+\mathbf{q},s'}) \cdot \boldsymbol{\epsilon}_{r,\mathbf{q}} \\ & + c_{\mathbf{p},s}^\dagger c_{\mathbf{p}+\mathbf{q},s'} a_{r\mathbf{q}}^\dagger (\bar{u}_{\mathbf{p},s} \boldsymbol{\gamma} u_{\mathbf{p}+\mathbf{q},s'}) \cdot \boldsymbol{\epsilon}_{r,\mathbf{q}}^* + c_{\mathbf{p},s}^\dagger d_{-\mathbf{p}-\mathbf{q},s'}^\dagger a_{r\mathbf{q}}^\dagger (\bar{u}_{\mathbf{p},s} \boldsymbol{\gamma} v_{-\mathbf{p}-\mathbf{q},s'}) \cdot \boldsymbol{\epsilon}_{r,\mathbf{q}}^* \\ & + d_{\mathbf{p},s} c_{-\mathbf{p}-\mathbf{q},s'} a_{r\mathbf{q}} (\bar{v}_{\mathbf{p},s} \boldsymbol{\gamma} u_{-\mathbf{p}-\mathbf{q},s'}) \cdot \boldsymbol{\epsilon}_{r,\mathbf{q}} + d_{\mathbf{p},s} d_{\mathbf{p}+\mathbf{q},s'}^\dagger a_{r\mathbf{q}} (\bar{v}_{\mathbf{p},s} \boldsymbol{\gamma} v_{\mathbf{p}+\mathbf{q},s'}) \cdot \boldsymbol{\epsilon}_{r,\mathbf{q}} \\ & + d_{\mathbf{p},s} c_{-\mathbf{p}+\mathbf{q},s'} a_{r\mathbf{q}}^\dagger (\bar{v}_{\mathbf{p},s} \boldsymbol{\gamma} u_{-\mathbf{p}+\mathbf{q},s'}) \cdot \boldsymbol{\epsilon}_{r,\mathbf{q}}^* + d_{\mathbf{p},s} d_{\mathbf{p}-\mathbf{q},s'}^\dagger a_{r\mathbf{q}}^\dagger (\bar{v}_{\mathbf{p},s} \boldsymbol{\gamma} v_{\mathbf{p}-\mathbf{q},s'}) \cdot \boldsymbol{\epsilon}_{r,\mathbf{q}}^* \end{aligned} \right\} \quad (4.67)$$

FIGURE 4.1

Les huit différents vertex associés à l'interaction électron-photon dans la théorie des perturbations ordinaire. Dans la théorie covariante des perturbations, c'est-à-dire dans les diagrammes de Feynman, tous ces vertex se fondent en un seul.



Chacun de ces huit termes correspond à un processus particulier et à un type de vertex distinct dans la théorie des perturbations ordinaires. En fonction de véritables diagrammes de Feynman, ces huit termes ne représentent qu'un seul vertex, soit le vertex fondamental de la QED. Les huit termes sont illustrés à la figure 4.1, dans le même ordre que ci-haut. Par convention et afin de les distinguer, les lignes associées aux électrons sont munies d'une flèche orientée de l'état initial vers l'état final, alors que les lignes associées aux positrons sont munies de flèches orientées en sens contraire. Dans les véritables diagrammes de Feynman, ces flèches sont conservées, mais on ne distingue pas les électrons des positrons virtuels; on se doit cependant de conserver la continuité des flèches aux vertex. Selon une interprétation commune, les antiparticules sont comme des particules qui se propagent à l'envers dans le temps (d'où le signe négatif de l'énergie dans l'exponentielle associée à la dépendance temporelle).

4.C.2 Règles de Feynman

La théorie covariante des perturbations, déjà évoquée à la section 2.C.4, s'applique également à l'électrodynamique. Les règles générales de Feynman sont les mêmes. Ce sont les expressions des lignes et des vertex qui sont différentes. Rappelons donc la totalité des règles, adaptées au cas présent :

1. Les règles suivantes visent à construire une expression mathématique pour $i\mathcal{M}$, \mathcal{M} étant l'amplitude invariante du processus dans la normalisation relativiste. Les facteurs entrant dans cette expression sont résumés dans le tableau 4.1 et décrits dans ce qui suit.
2. Pour un processus donné, identifier les états initial et final et dessiner les lignes correspondantes. Les lignes sont affublées de flèches pour distinguer les particules (flèches dirigées vers le haut) des antiparticules (flèches dirigées vers le bas). Un photon est représenté par une ligne ondulée, sans flèche (le photon est sa propre antiparticule). Ensuite on doit inscrire un facteur associé à chaque ligne externe, comme indiqué dans les quatre premiers éléments du tableau 4.1 : à un fermion initial, on associe un spineur $u(\mathbf{p}, s)$ associé à l'impulsion \mathbf{p} et à la polarisation s . Pour un fermion final, c'est le spineur conjugué \bar{u} qui apparaît. Pour un antifermion, u est remplacé par \bar{v} et \bar{u} par v . Pour un photon initial, on fait intervenir le vecteur de polarisation ε_μ , et son conjugué pour un photon final. Ce quadrivecteur polarisation dépend du vecteur d'onde et de la jauge utilisée. Par exemple, dans la jauge de Coulomb,

$$\varepsilon_0 = 0 \quad \text{et} \quad \boldsymbol{\varepsilon}(\mathbf{q}) \cdot \mathbf{q} = 0 \quad (4.68)$$

3. Construire les diagrammes possibles à l'ordre N de la théorie des perturbations en introduisant N vertex, auxquels sont reliés les lignes externes et des lignes internes (correspondant à des particules virtuelles) introduites au besoin.
4. Chaque ligne porte une certaine quadri-impulsion : notons-la p_i pour les lignes externes, et q_i pour les lignes internes. La direction de cette quadri-impulsion est arbitraire, c'est-à-dire affaire de convention. On peut faire une analogie avec les lois de Kirchhoff dans la théorie des circuits. La quadri-impulsion est conservée à chaque vertex.
5. À chaque vertex, insérer un facteur $-ie(\gamma^\mu)_{\beta\alpha}$. L'indice μ est attaché à la ligne de photon reliée au vertex, l'indice β à la ligne de fermion sortante et α à la ligne de fermion entrante. Ajouter un facteur $(2\pi)^4\delta(k_1+k_2+k_3)$, où $k_{1,2,3}$ sont les quadri-impulsions *entrantes* (changer le signe s'il s'agit de quadri-impulsions sortantes). Important : chaque vertex comporte deux indices de Dirac et un indice d'espace-temps.
6. À chaque ligne interne est associé un *propagateur*. Pour un photon virtuel, ce propagateur est

$$-i\frac{g_{\mu\nu}}{q^2}$$

où q est la 4-impulsion du photon virtuel. Les indices μ et ν sont ceux qui sont attachés aux deux vertex reliés par la ligne interne. Pour un fermion, ce propagateur est

$$i\frac{(p_\mu\gamma^\mu + m)_{\beta\alpha}}{p^2 - m^2} = i\frac{(\not{p} + m)_{\beta\alpha}}{p^2 - m^2}$$

où les indices de Dirac β et α sont ceux qui sont attachés aux deux vertex reliés par la ligne interne, du côté de la ligne interne (car chaque vertex comporte deux indices de Dirac, un associé à chaque ligne qui sort du vertex).

7. Intégrer sur toute quadri-impulsion interne q_i :

$$\int \frac{d^4 q_i}{(2\pi)^4} \dots$$

En pratique, si le diagramme ne contient pas de boucle fermée, les fonctions deltas qui imposent la conservation de la 4-impulsion à chaque vertex rendent ces intégrales triviales. Chaque boucle du diagramme correspond à une réelle intégration. Ces intégrales constituent la principale difficulté calculatoire, notamment en raison de leur caractère singulier.

8. La conservation de la quadri-impulsion entre les états initial et final est une conséquence de sa conservation à chaque vertex. Cependant, le facteur global $(2\pi)^4\delta(p_1 + p_2 + \dots - p_n)$ associé à la conservation de l'énergie-impulsion doit être amputé, pour obtenir l'amplitude \mathcal{M} (par opposition à M). Voir à cet effet la discussion de la section 1.E.2).
9. Les diagrammes qui diffèrent par la permutation de deux lignes externes associées à des fermions indiscernables doivent être combinés avec un signe relatif. De plus, toute boucle interne de fermions doit être accompagnée d'un facteur -1 .

Ces règles sont applicables dans la normalisation relativiste des états. Dans cette normalisation, les facteurs $(2E_p)^{-1/2}$ qui apparaissent devant les spineurs (3.105, 3.107) disparaissent. Les spineurs

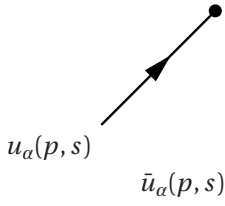
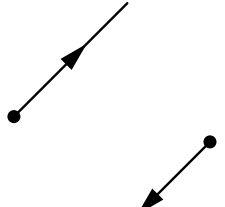
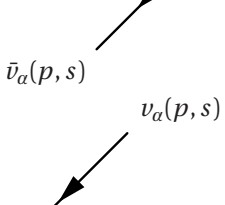
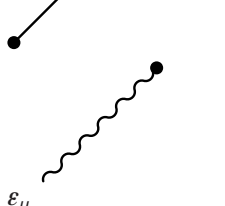
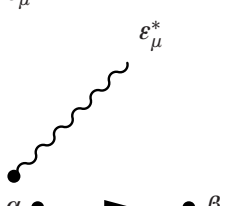
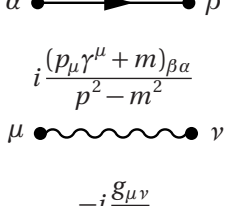
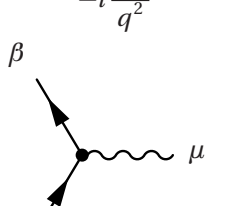
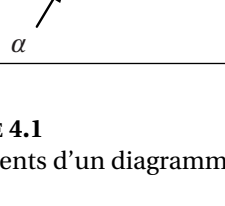

fermion initial	
fermion final	
antifermion initial	
antifermion final	
photon initial	
photon final	
fermion virtuel	
photon virtuel	
vertex	

TABLE 4.1
Facteurs associés aux différents éléments d'un diagramme de Feynman.

associés aux ondes planes sont donc plutôt les suivants :

$$u_{\mathbf{p},1} = \frac{1}{\sqrt{E_{\mathbf{p}} + m}} \begin{pmatrix} E_{\mathbf{p}} + m \\ 0 \\ p_z \\ p_x + i p_y \end{pmatrix} \quad u_{\mathbf{p},2} = \frac{1}{\sqrt{E_{\mathbf{p}} + m}} \begin{pmatrix} 0 \\ E_{\mathbf{p}} + m \\ p_x - i p_y \\ -p_z \end{pmatrix} \quad (4.69)$$

$$v_{\mathbf{p},1} = \frac{1}{\sqrt{E_{\mathbf{p}} + m}} \begin{pmatrix} p_x - i p_y \\ -p_z \\ 0 \\ E_{\mathbf{p}} + m \end{pmatrix} \quad v_{\mathbf{p},2} = \frac{1}{\sqrt{E_{\mathbf{p}} + m}} \begin{pmatrix} p_z \\ p_x + i p_y \\ E_{\mathbf{p}} + m \\ 0 \end{pmatrix} \quad (4.70)$$

Ces spineurs ont les propriétés suivantes :

$$\bar{u}_{\mathbf{p},s} u_{\mathbf{p},s'} = 2m \delta_{ss'} \quad \bar{v}_{\mathbf{p},s} v_{\mathbf{p},s'} = -2m \delta_{ss'} \quad \bar{u}_{\mathbf{p},s} v_{\mathbf{p},s'} = \bar{v}_{\mathbf{p},s} u_{\mathbf{p},s'} = 0 \quad (4.71)$$

Les spineurs forment aussi une base complète dans l'espace des spineurs, ce qui se traduit par les relations

$$\sum_{s=1,2} u_{\alpha}(\mathbf{p}, s) \bar{u}_{\beta}(\mathbf{p}, s) = (\not{\mathbf{p}} + m)_{\alpha\beta} \quad \sum_{s=1,2} v_{\alpha}(\mathbf{p}, s) \bar{v}_{\beta}(\mathbf{p}, s) = (\not{\mathbf{p}} - m)_{\alpha\beta} \quad (4.72)$$

où nous avons introduit la notation abrégée

$$\not{p} \equiv p_{\mu} \gamma^{\mu} \quad \not{\partial} \equiv \partial_{\mu} \gamma^{\mu} \quad \not{A} \equiv A_{\mu} \gamma^{\mu} \quad (4.73)$$

dans laquelle une barre oblique traversant un symbole signifie la contraction avec les matrices de Dirac. Cette abréviation est couramment utilisée dans le domaine. Nous avons aussi mis les indices de vecteur d'onde et de polarisation entre parenthèses, afin de laisser de la place aux indices de spineur. Ceci peut se faire selon le contexte.

Somme sur les polarisations et formules des traces

Très souvent, on doit faire la somme de la section efficace sur les polarisations possibles de l'état final (si la polarisation des particules finales n'est pas observée dans une expérience) et la moyenne sur les polarisations des particules initiales (si les particules initiales ne sont pas préparées dans un état de polarisation donnée). Il s'agit d'une somme des carrés des amplitudes, et elle permet généralement de simplifier considérablement les résultats.

La formule suivante, dite *formule de Casimir*, permet d'effectuer cette somme :

$$\sum_{s_1, s_2} [\bar{u}(p_1, s_1) \Gamma u(p_2, s_2)] [\bar{u}(p_1, s_1) \Gamma' u(p_2, s_2)]^* = \text{tr} [\Gamma (\not{p}_2 + m_2) \bar{\Gamma}' (\not{p}_1 + m_1)] \quad (4.74)$$

où Γ et Γ' sont des matrices 4×4 , comme des matrices de Dirac ou des combinaisons ou produits de matrices de Dirac. Nous avons introduit la notation

$$\bar{\Gamma} = \gamma^0 \Gamma^{\dagger} \gamma^0 \quad (4.75)$$

Si $\Gamma = \gamma^\mu$, alors $\bar{\Gamma} = \Gamma$.

La démonstration de la formule de Casimir est immédiate à l'aide des relations de complétude (4.72). Notons d'abord que

$$\begin{aligned} [\bar{u}(p_1, s_1) \Gamma' u(p_2, s_2)]^* &= [u^\dagger(p_1, s_1) \gamma^0 \Gamma' u(p_2, s_2)]^\dagger \\ &= u^\dagger(p_2, s_2) (\Gamma')^\dagger \gamma^0 u(p_1, s_1) \\ &= u^\dagger(p_2, s_2) \gamma^0 \gamma^0 (\Gamma')^\dagger \gamma^0 u(p_1, s_1) \quad ((\gamma^0)^2 = 1) \\ &= \bar{u}(p_2, s_2) \bar{\Gamma}' u(p_1, s_1) \end{aligned} \quad (4.76)$$

Ensuite, détaillons les indices de spineurs :

$$\begin{aligned} \sum_{s_1, s_2} [\bar{u}(p_1, s_1) \Gamma u(p_2, s_2)] [\bar{u}(p_2, s_2) \bar{\Gamma}' u(p_1, s_1)] &= \sum_{s_1, s_2} \bar{u}_\alpha(p_1, s_1) \Gamma_{\alpha\beta} u_\beta(p_2, s_2) \bar{u}_\gamma(p_2, s_2) \bar{\Gamma}'_{\gamma\delta} u_\delta(p_1, s_1) \\ &= (\not{p}_1 + m)_{\delta\alpha} (\not{p}_2 + m)_{\beta\gamma} \Gamma_{\alpha\beta} \bar{\Gamma}'_{\gamma\delta} \\ &= \text{tr} [\Gamma (\not{p}_2 + m_2) \bar{\Gamma}' (\not{p}_1 + m_1)] \end{aligned} \quad (4.77)$$

Si on a affaire à des antifermions plutôt qu'à des fermions, on obtient $\not{p} - m$ plutôt que $\not{p} + m$.

Le calcul des traces s'effectue en utilisant les formules suivantes :

$\text{tr}(\gamma^\mu \gamma^\nu) = 4g^{\mu\nu} \quad \text{tr}(\gamma^\mu \gamma^\nu \gamma^\rho \gamma^\lambda) = 4(g^{\mu\nu} g^{\rho\lambda} - g^{\mu\rho} g^{\nu\lambda} + g^{\mu\lambda} g^{\rho\nu})$

(4.78)

La trace d'un produit d'un nombre impair de matrices de Dirac est nulle.

4.C.3 Exemples de processus

Indiquons quelques processus décrits par la théorie de l'interaction électron-photon, et les diagrammes de Feynman associés à l'ordre le plus bas de la théorie des perturbations.

Diffusion électron-muon

(4.79)

Dans ce processus, l'important est que les deux fermions qui diffusent l'un sur l'autre ne soient pas identiques. Les lignes sont étiquetées par le symbole représentant la particule quand il y a risque de confusion. Une application des règles de Feynman mène à l'amplitude suivante :

$$\mathcal{M} = \frac{4\pi\alpha}{(p_1 - p_3)^2} g_{\mu\nu} [\bar{u}_{p_3, s_3} \gamma^\mu u_{p_1, s_1}] [\bar{u}_{p_4, s_4} \gamma^\nu u_{p_2, s_2}] \quad \left(\alpha = \frac{e^2}{4\pi} \right) \quad (4.80)$$

La somme sur les polarisations (s_1, s_2, s_3, s_4) du carré de l'amplitude (divisée par 4 pour effectuer une moyenne sur les polarisations initiales), effectuée à l'aide de la formule de Casimir et des formules de traces, donne

$$\langle |\mathcal{M}|^2 \rangle = \frac{8(4\pi\alpha)^2}{(p_1 - p_3)^4} [(p_1 \cdot p_2)(p_3 \cdot p_4) + (p_1 \cdot p_4)(p_2 \cdot p_3) - m^2(p_1 \cdot p_3) - M^2(p_2 \cdot p_4) + 2m^2 M^2] \quad (4.81)$$

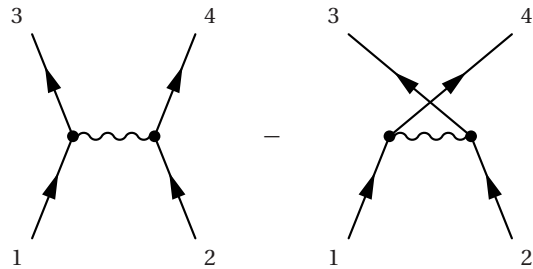
où m est la masse de l'électron et M celle du muon.

À partir de cette formule, on peut démontrer la formule de Mott pour la section différentielle de diffusion dans le repère du muon initial, en supposant que celui-ci est infiniment massif par rapport à l'électron :

$$\left\langle \frac{d\sigma}{d\Omega} \right\rangle = \left(\frac{\alpha}{2\mathbf{p}^2 \sin^2(\theta/2)} \right)^2 (m^2 + \mathbf{p}^2 \cos^2(\theta/2)) = \left(\frac{\alpha E}{2\mathbf{p}^2 \sin^2 \theta/2} \right)^2 (1 - \beta^2 \sin^2(\theta/2)) \quad (4.82)$$

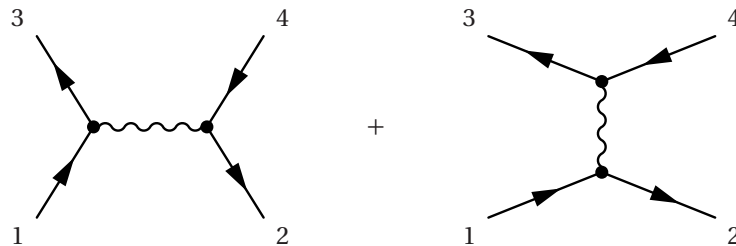
où \mathbf{p} et E sont la quantité de mouvement et l'énergie du projectile, θ est l'angle de diffusion, β est la vitesse du projectile.

Diffusion électron-électron ou Diffusion Møller


(4.83)

Dans ce cas, l'identité des deux particules en cause nous force à considérer deux diagrammes qui diffèrent par une permutation des lignes sortantes.

Diffusion électron-positron ou Diffusion Bhabha


(4.84)

Dans le référentiel du centre de masse, la section différentielle de diffusion pour des électrons ultrarelativistes ($m^2 \ll s$) est

$$\frac{d\sigma}{d\Omega} = \frac{\alpha^2}{2s} \left(u^2 \left(\frac{1}{s} + \frac{1}{t} \right)^2 + \left(\frac{t}{s} \right)^2 + \left(\frac{s}{t} \right)^2 \right) \quad (4.85)$$

où s, t, u sont les variables de Mandelstam et θ l'angle de diffusion dans ce repère.

Production de paires de particules chargées

Le deuxième diagramme associé à la diffusion Bhabba peut aussi servir au cas où les particules produites ne sont pas une paire électron-positron, mais plutôt une paire muon-antimuon, ou quark-antiquark. Dans ce cas le diagramme de gauche ne s'applique pas et la section différentielle, dans le référentiel du centre de masse, est

$$\frac{d\sigma}{d\Omega} = \frac{\alpha^2}{4s} \beta (1 + \cos^2 \theta + (1 - \beta^2) \sin^2 \theta) N_c Q_f^2 \quad (4.86)$$

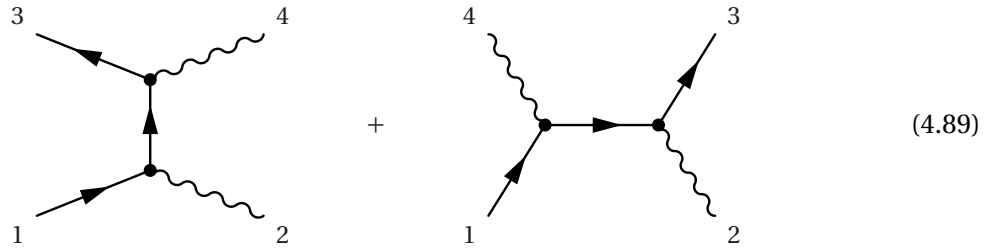
où θ est l'angle de diffusion dans ce repère, β est la vitesse des particules et s la première variable de Mandelstam. N_c est égal à 3 pour des quarks, et à 1 pour des muons (ce facteur est le nombre de couleurs, voir p. 165). Q_f est la charge électrique de la particule produite, en unités de la charge élémentaire (1 pour un muon, $\frac{2}{3}$ pour un quark u , etc.). Dans la limite ultrarelativiste, cette expression devient

$$\frac{d\sigma}{d\Omega} = \frac{\alpha^2}{4s} (1 + \cos^2 \theta) N_c Q_f^2 \quad (4.87)$$

et la section efficace associée est

$$\sigma = N_c Q_f^2 \frac{4\pi}{3} \frac{\alpha^2}{s} \quad (4.88)$$

Diffusion électron-photon ou Diffusion Compton



Notons ici que deux diagrammes sont nécessaires, différant par une permutation des photons initial et final. L'amplitude obtenue à l'aide de ces deux diagrammes mène à la section différentielle de Klein-Nishina, exprimé ici dans le repère du laboratoire (l'électron initial est au repos) :

$$\frac{d\sigma}{d\Omega} = \frac{1}{2} \alpha^2 r_c^2 \left(\frac{E'}{E} \right)^2 \left(\frac{E'}{E} + \frac{E}{E'} - 1 - \cos^2 \theta \right) \quad (4.90)$$

où θ est l'angle de diffusion et E, E' sont les énergies du photon initial et final, respectivement. On a la relation bien connue

$$\frac{1}{E'} = \frac{1}{E} + \frac{1}{m} (1 - \cos \theta) \quad (4.91)$$

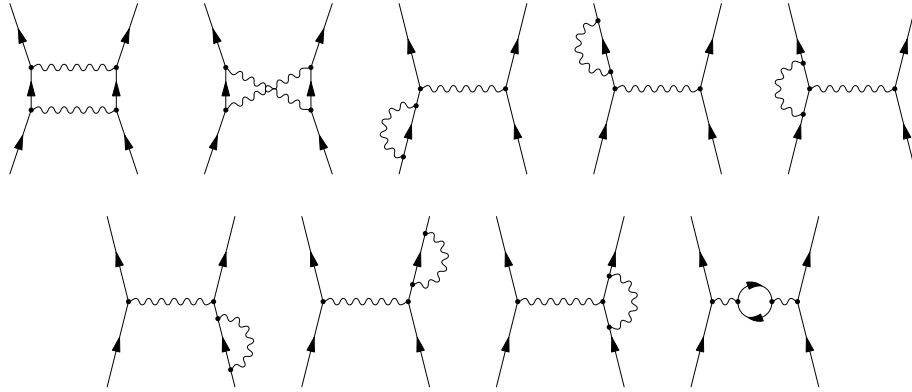


FIGURE 4.2

Diagrammes contribuant à l'amplitude de diffusion électron-électron, à l'ordre 4 de la théorie des perturbations. À ces diagrammes on doit ajouter ceux dont les deux pattes supérieures sont échangées, avec une phase relative de -1 .

D

Corrections radiatives et renormalisation

4.D.1 Corrections d'ordre supérieur

Pour un processus donné, les diagrammes à l'ordre le plus bas de la théorie des perturbations, comme ceux illustrés ci-haut, ne comportent aucune boucle, c'est-à-dire qu'ils ont la topologie d'un arbre. Ces diagrammes sont tous relativement simples à calculer, en dépit d'un certain nombre de traces de matrices de Dirac à effectuer.

Pour un processus donné, par exemple la diffusion électron-électron, les diagrammes contribuant à l'ordre suivant comportent tous une boucle, comme indiqué à la figure 4.2. On montre que la présence d'une boucle entraîne l'existence d'une quadri-impulsion libre, non déterminée par la conservation de la quadri-impulsion à chaque vertex. En fait, il est assez clair que la quadri-impulsion circulant dans une boucle insérée dans une ligne externe ou interne peut prendre une valeur arbitraire sans affecter la conservation de la quadri-impulsion le long de cette ligne. Le problème est que l'intégrale correspondante sur la quadri-impulsion est souvent divergente.

Ces divergences dans la théorie des perturbations sont apparues très tôt dans la théorie – avant que cette dernière soit formulée à l'aide de diagrammes – et en ont retardé le développement pendant presque vingt ans. Leur origine tient au traitement continu de l'espace-temps : ce continuum entraîne que les vecteurs d'onde prennent toutes les valeurs jusqu'à l'infini et les divergences proviennent du comportement des intégrales sur les impulsions quand elles tendent vers l'infini – on dit que ce sont des divergences *ultraviolettes*. Pour donner un sens aux diagrammes divergents, il faut premièrement introduire une *régularisation*, c'est-à-dire une procédure cinématique qui rende le diagramme fini. Plusieurs procédures différentes existent, mais elles ont en commun d'introduire une échelle d'énergie Λ , dite *coupure*, qui représente à peu près la valeur maximale du vecteur d'onde.

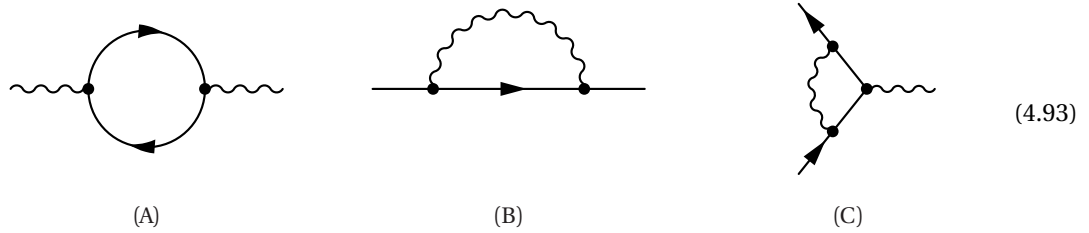
On dit qu'un diagramme D possède un *degré de divergence* ζ s'il se comporte comme Λ^ζ . La présence d'une divergence ne résulte pas automatiquement de celle d'une boucle dans le diagramme. Elle dépend aussi de sa topologie, du nombre de pattes externes et de leur type (photon ou électron). Par analyse dimensionnelle, on montre facilement que ζ devrait avoir la valeur suivante, dit *degré superficiel de divergence* :

$$\zeta_{\text{sup.}}(D) = 4 - \frac{3}{2}F - B \quad (4.92)$$

où F est le nombre de pattes (lignes externes) représentant des fermions et B le nombre de pattes représentant des bosons. En électrodynamique quantique, l'invariance de jauge complique sensiblement ce calcul et le degré de divergence réel d'un diagramme est parfois inférieur à $\zeta_{\text{sup.}}$. En somme,

- Si $\zeta = 0$, le diagramme diverge comme le logarithme de la coupure Λ .
- Si $\zeta > 0$, le diagramme diverge comme Λ^ζ .
- Si $\zeta < 0$, le diagramme peut être convergent (mais des sous-diagrammes peuvent être divergents).

En électrodynamique quantique, seuls trois diagrammes ou sous-diagrammes sont divergents, illustrés ci-dessous :



1. Le diagramme (A) contribue à la polarisation du vide. Il représente l'effet sur la propagation de la lumière de la présence de paires électron-positron virtuelle dans le vide.
2. Le diagramme (B) contribue à l'énergie propre (self-énergie) de l'électron. Il représente l'effet sur un électron de son propre champ électromagnétique.
3. Le diagramme (C) porte le nom de *correction de vertex* et contribue à redéfinir la charge de l'électron et son interaction avec le champ électromagnétique en général.

C'est le travail de Bethe, Feynman, Schwinger et Tomonaga à la fin des années 1940 qui a permis de contourner la difficulté posée par les diagrammes divergents (et avant eux, Stückelberg, sans impact sur la communauté scientifique, malheureusement). Une procédure, appelée *renormalisation*, a été élaborée afin d'absorber des quantités formellement infinies – mais en pratique inobservables – dans une redéfinition de la charge e , des masses des particules et de la normalisation des champs (d'où le nom de renormalisation). Cette procédure est essentielle afin de donner un sens aux ordres plus élevés de la théorie des perturbations et permet dans quelques cas d'effectuer des calculs d'une remarquable précision. Les corrections aux ordres supérieurs de la théorie des perturbations portent généralement le nom de *corrections radiatives*.

Déplacement de Lamb Le phénomène physique qui a initialement stimulé la technique de renormalisation est le déplacement de Lamb. Il s'agit d'une petite différence d'énergie entre les niveaux $2S_{1/2}$ ($n = 2, l = 0, j = \frac{1}{2}$) et $2P_{1/2}$ ($n = 2, l = 1, j = \frac{1}{2}$) de l'atome d'hydrogène. Ces deux niveaux sont dégénérés selon la solution de l'équation de Dirac (l'énergie ne dépend que de n et de j , voir éq. (4.38)). Par contre, une différence d'énergie d'environ 1 000 MHz est observée entre les deux niveaux.² Cette différence a été expliquée par un effet des corrections radiatives en 1947 par H. Bethe.³ On montre que le déplacement de Lamb est donné par l'expression suivante :

2. W.E. Lamb et R.C. Retherford, *Phys. Rev.* **72**, 241 (1947).

3. H. Bethe, *Phys. Rev.* **72** 339 (1947).

$$\begin{aligned}\Delta E_{\text{Lamb}}(n, 0, \tfrac{1}{2}) &= \alpha^5 m c^2 \frac{1}{4n^3} k(n, 0) & (l = 0) \\ \Delta E_{\text{Lamb}}(n, l, j) &= \alpha^5 m c^2 \frac{1}{4n^3} \left(k(n, l) \pm \frac{1}{\pi(j + 1/2)(l + 1/2)} \right) & (l \neq 0)\end{aligned}\quad (4.94)$$

où $k(n, l)$ est un facteur numérique qui dépend de n et l ; ce facteur est d'ordre ~ 13 dans le cas $l = 0$, mais beaucoup plus petit ($< 0,05$) dans le cas $l \neq 0$. La valeur de j est donnée par $l \pm 1/2$, où le signe \pm est celui qui apparaît dans la deuxième équation ci-dessus. La différence d'énergie entre les niveaux $2S_{1/2}$ et $2P_{1/2}$ est alors

$$\Delta E_{\text{Lamb}}(2, 1, \tfrac{1}{2}) - \Delta E_{\text{Lamb}}(2, 0, \tfrac{1}{2}) \quad (4.95)$$

Moment magnétique de l'électron

L'exemple le plus remarquable de correction radiative est le calcul du moment magnétique anormal de l'électron, le célèbre facteur g . Rappelons que le moment magnétique de l'électron est de grandeur $g\mu_B$, où $\mu_B = e/2m_e$ est le magnéton de Bohr et g le facteur de Landé, communément appelé « facteur g ». Si l'électron se comportait comme une distribution de charge classique tournant sur elle-même, comme une microscopique boule de billard, son facteur g serait l'unité. On montre que l'équation de Dirac mène à la valeur $g = 2$, ce qui constitue l'un des succès de cette équation. Par contre, la valeur mesurée de g est différente de 2, même si elle en est très proche. La valeur expérimentale acceptée, provenant d'expériences de résonances très précises, est la suivante :⁴

$$g/2 = 1,001\,159\,652\,180\,73\,(28) \quad (4.96)$$

L'électrodynamique quantique permet de calculer la déviation de g par rapport à 2. Cette déviation provient exclusivement de corrections radiatives. Le premier calcul en ce sens a été effectué par Schwinger en 1948⁵, au deuxième ordre en théorie des perturbations, avec le résultat suivant :

$$\frac{g}{2} = 1 + \frac{\alpha}{2\pi} = 1.001\,161\,410\dots \quad (4.97)$$

Le calcul depuis a été mené jusqu'à l'ordre 10 en théorie des perturbations – correspondant à cinq boucles par diagramme. L'ordre 10 à lui seul comporte 12 672 diagrammes, presque tous générés par un programme et calculés par des méthodes numériques. Le résultat s'exprime en fonction de la constante de structure fine, dont la valeur la plus précise, obtenue par une méthode indépendante du présent calcul, est⁶

$$\alpha^{-1} = 137,035\,999\,049(90) \quad (4.98)$$

La valeur de $g/2$ prédite par la QED est alors⁷

$$g/2 = 1,001\,159\,652\,181\,78\,(6)(4)(2)(77) \quad (4.99)$$

où les incertitudes sur les derniers chiffres entre parenthèses proviennent de divers facteurs : calculs numériques d'intégrales, contributions des hadrons et incertitude sur la valeur de α , cette dernière constituant la source d'erreur la plus importante. La différence entre les valeurs mesurée et calculée

4. D. Hanneke et al., Phys. Rev. Lett. **100**, 120801 (2008); Phys. Rev. A **83**, 052122 (2012).

5. J. Schwinger, Phys. Rev. **73**, 416L (1948).

6. R. Bouchendira et al., Phys. Rev. Lett. **106**, 080801 (2011).

7. Aoyama et al., Phys. Rev. Lett. **109**, 111807 (2012).

de $g/2$ est compatible avec zéro dans les barres d'erreur et est d'ordre 10^{-12} . Ceci est souvent cité comme l'accord théorie-expérience le plus précis de toute la science.

Comme l'erreur la plus importante sur la prédiction théorique de $g/2$ provient de la valeur même de α , il est raisonnable de redéfinir la valeur acceptée de α à partir de cet accord, ce qui donne

$$\alpha^{-1} = 137,035\,999\,172\,7(68)(46)(19)(331) \quad (4.100)$$

4.D.2 Constante de couplage variable

En tenant compte des corrections radiatives, on peut effectivement remplacer le vertex simple, donné par $-ie(\gamma^\mu)_{\beta\alpha}$ dans les règles de Feynman, par une fonction plus complexe, qu'on notera $-ie\Gamma_{\beta\alpha}^\mu(p, p')$ et qui dépend des quadrivecteurs des fermions entrant (p) et sortant (p'). Cette quantité, appelée « fonction de vertex », contient les corrections radiatives à l'interaction entre les fermions chargés et le champ électromagnétique :

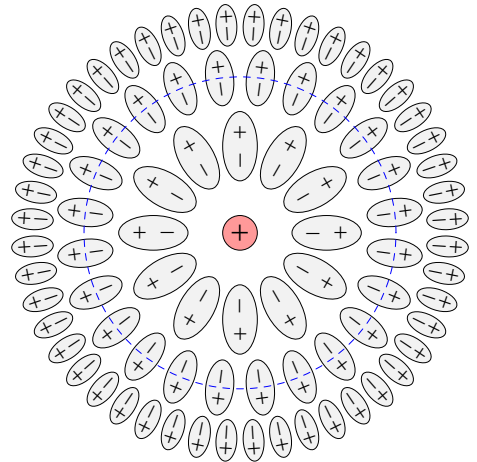
$$-ie\Gamma = \text{diagram 1} + \text{diagram 2} + \text{diagram 3} + \text{diagram 4} + \dots \quad (4.101)$$

Cette fonction dépend des impulsions entrante et sortante, dont deux sont indépendantes (la troisième, l'impulsion du photon entrant q , est donnée par $p' - p$). On peut montrer qu'elle a la structure suivante :

$$\Gamma^\mu(p, p') = \gamma^\mu(1 + F_1(q^2)) + \frac{i}{2m}\sigma^{\mu\nu}q_\nu F_2(q^2) \quad \text{où} \quad \sigma^{\mu\nu} := \frac{i}{2}[\gamma^\mu, \gamma^\nu] \quad (4.102)$$

Ici $q = p' - p$ et F_1 et F_2 sont des fonctions scalaires appelées « facteurs de forme ». On montre que le moment magnétique de l'électron est proportionnel à $F_2(0)$.

FIGURE 4.3
L'effet écran.



Bien sûr l'évaluation précise des fonctions F_1 et F_2 est complexe, et demande un traitement des divergences par la procédure de renormalisation. Cependant, l'important ici n'est pas le calcul de ces

fonctions, mais le fait qu'on pourrait, les connaissant en principe, obtenir l'interaction effective des électrons avec le champ électromagnétique, et que cette interaction dépend de l'échelle d'énergie q^2 du photon virtuel impliqué (notons qu'en général les impulsions p , p' et q ne sont pas sur la couche de masse). Cela mène à la notion de « constante de couplage variable », par laquelle on peut définir une charge effective $e(q^2) = e(1 + F_1(q^2))$ qui varie en fonction de l'échelle d'énergie. Or, en raison des divergences associées au continuum, on ne peut vraiment bien définir ce facteur F_1 que par rapport à un « point de soustraction », ou « point de référence ». Autrement dit, on ne peut pas connaître F_1 (ou F_2) en absolu, mais seulement par rapport à une référence choisie de q^2 . La constante structure fine $\alpha = e^2/4\pi$ n'est donc pas une véritable constante, mais une fonction $\alpha(q^2)$ qui varie légèrement en fonction de l'échelle d'énergie q^2 associée au vertex. On montre que, dans une certaine approximation (diagrammes ne comportant qu'une boucle),

$$\frac{1}{\alpha(q^2)} = \frac{1}{\alpha(q_0^2)} - 4N_g \ln \frac{q^2}{q_0^2} \quad (4.103)$$

où $N_g = 3$ est le nombre de familles de particules élémentaires et où l'effet tient ici compte de toutes les particules chargées. Cela signifie que, quand q^2 augmente, $\alpha(q^2)$ augmente aussi, mais faiblement, en raison du logarithme. Cette augmentation de la constante de couplage avec q^2 peut être comprise par la notion d'écrantage : dans un milieu polarisable, la charge effective d'une source est de plus en plus grande lorsqu'on l'observe de plus en plus proche, car la polarisation du milieu entourant la charge tend à la masquer partiellement (voir 4.3). Le vide aussi est polarisable, car des paires virtuelles électron-positron peuvent être créées, et donc la charge élémentaire, ou la constante de structure fine $\alpha = e^2/4\pi$, augmente avec la résolution spatiale, c'est-à-dire augmente avec l'échelle d'énergie associée au processus.

4.D.3 Interprétation de la renormalisation

La procédure de renormalisation suggérée à la fin des années 1940, bien qu'elle permette des calculs précis et non équivoques, a longtemps suscité la méfiance, en raison de son caractère mathématique peu rigoureux. Elle revient, selon plusieurs à l'époque, à soustraire des quantités infinies l'une de l'autre, ou encore à effectuer un développement perturbatif en puissances d'une quantité infinie. Depuis les années 1980, notre interprétation de la renormalisation est beaucoup plus concrète et pose moins de problèmes, en partie suite aux travaux de Kenneth WILSON dans les années 1970 sur le groupe de renormalisation. Il est maintenant généralement accepté que la théorie des champs n'est pas une théorie ultime, mais plutôt une théorie effective, applicable à des échelles de longueur beaucoup plus grandes qu'une certaine échelle fondamentale, notée Λ^{-1} (telle la longueur de Planck, par exemple). Les prédictions d'une théorie des champs dépendent donc de ses paramètres (par exemple la masse de l'électron et la constante de structure fine α), ainsi que de cette échelle Λ . Cependant, les prédictions de cette théorie ne doivent pas dépendre de Λ , si Λ est très grand par rapport aux échelles d'énergie étudiées. En fait, il existe une faible dépendance (dite logarithmique, car elle fait intervenir $\log \Lambda$) des prédictions sur Λ . Cette dépendance peut être absorbée comme une dépendance des paramètres de la théorie sur Λ . Ainsi, la constante de structure fine deviendrait une fonction légèrement dépendante de Λ : $\alpha(\Lambda)$. Une théorie dont les paramètres dépendent faiblement de Λ et sont en nombre fini est dite *renormalisable*. L'électrodynamique quantique, et plus généralement les théories de jauge, sont renormalisables. De telles théories ont en principe un réel pouvoir

de prédiction. Il est cependant facile de concevoir des termes qu'on pourrait ajouter au lagrangien d'une théorie des champs, qui ne sont pas renormalisables, c'est-à-dire qui diminuent rapidement – en loi de puissance au lieu de logarithmiquement – quand Λ augmente. Ces termes sont absents du modèle standard, non par exigence *a priori*, comme on le pensait autrefois, mais bien parce qu'en réalité l'échelle Λ est très grande, de sorte que si ces termes avaient une valeur « normale » pour des phénomènes se déroulant à des échelles d'énergie de l'ordre de Λ , ils sont complètement négligeables aux échelles d'énergie accessibles aux accélérateurs actuels. Une théorie des champs contenant de tels termes à notre échelle n'aurait pas beaucoup de sens, car cela supposerait que les termes correspondants du hamiltonien fondamental défini à l'échelle Λ sont incommensurablement plus grands que ceux qui donnent naissance aux termes renormalisables. Il faudrait aussi que de tels termes soient en nombre infini, ce qui enlèverait tout pouvoir de prédiction à la théorie.

Par contre, un terme comme la masse de l'électron dépend aussi fortement de l'échelle Λ , mais dans l'autre sens : il est pratiquement proportionnel à Λ . En fait, tout repose sur la dimension (au sens des unités) de chaque paramètre. Les paramètres du lagrangien ayant les unités d'une puissance positive de la masse sont en nombre fini (habituellement, seule la masse elle-même est dans ce cas) et ne posent pas de problème en raison de ce nombre fini. Les paramètres ayant les unités d'une puissance négative de la masse sont en nombre potentiellement infini et sont justement *non renormalisables*. Enfin, les paramètres ayant une dimension nulle (ou sans unité, telle la constante de structure fine α) ne sont sujets qu'à des variations logarithmiques en fonction de Λ . Plus précisément, un paramètre λ ayant les unités d'une masse à la puissance h , devrait avoir l'expression suivante dans une théorie ultime : $\lambda(\Lambda) = \lambda_0 \Lambda^h$, où λ_0 est une constante sans unité, ou dépendant logarithmiquement de Λ . Si $h < 0$ et que Λ est énorme en comparaison des échelles d'énergie étudiées – par exemple en comparaison de l'énergie des projectiles dans un processus de collision dans le repère du centre de masse – alors il faut que λ soit très petit, trop petit pour avoir des conséquences observables.

Comment détermine-t-on les dimensions des différents termes? En se basant sur le fait que l'action est sans unité. Donc la densité lagrangienne a les unités L^{-4} (en quatre dimensions d'espace-temps). Il s'ensuit que le champ de Dirac a les unités $[\psi] = L^{-3/2} = M^{3/2}$. Par exemple, un terme d'interaction formé de quatre spineurs de Dirac (ne correspondant pas nécessairement à quatre espèces de particules identiques) :

$$\lambda \bar{\psi}_1 \psi_2 \bar{\psi}_3 \psi_4 \quad (4.104)$$

est tel que la constante λ a les unités $L^2 = M^{-2}$. Un tel terme est non renormalisable. S'il est présent dans une théorie (comme la théorie de Fermi des interactions faibles), c'est qu'il est à la fois faible (c'est le cas) et qu'il provient d'une théorie plus fondamentale à une échelle d'énergie grande, mais accessible (c'est aussi le cas : la théorie électrofaible du modèle standard, caractérisée par une échelle d'énergie de l'ordre de 100 GeV, soit la masse des W^\pm et du Z_0).

Ainsi, il n'y a pas de différence de philosophie fondamentale entre une théorie des champs décrivant les particules élémentaires et une théorie des champs décrivant le comportement aux grandes longueurs d'onde des électrons dans un solide cristallin. Dans ce cas, il est bien connu que la théorie repose sur une théorie plus fondamentale faisant intervenir le réseau cristallin et des quantités discrètes dans cette théorie sont remplacées par des quantités continues dans la théorie des champs (approximation du continuum). Par exemple, un terme impliquant une différence entre deux sites du cristal serait remplacé par un développement de Taylor, dont seul le premier terme non trivial

(la première dérivée, par exemple) serait conservé. Les termes impliquant les dérivées secondes seraient alors « non renormalisables » et laissés de côté, à moins qu'on s'intéresse à des échelles de longueur qui ne sont pas trop éloignées de l'échelle fondamentale Λ (dans ce cas, le pas de réseau inverse du cristal).

E Couplage de Yukawa et électrodynamique scalaire

4.E.1 Interaction de Yukawa

Dans les chapitres précédents, nous avons défini les propriétés du champ scalaire ϕ , qui représente un boson de spin zéro, et celles du champ de Dirac ψ , qui représente un fermion de spin $\frac{1}{2}$. L'interaction de ces deux types de particules peut être décrite par un lagrangien d'interaction invariant de Lorentz dont la forme la plus simple est la suivante :

$$\mathcal{L}_I = -g\bar{\psi}\psi\phi \quad (4.105)$$

Prob. 4.7 Ce type d'interaction porte le nom de *couplage de Yukawa*, en raison de son rôle dans la théorie de Yukawa des interactions fortes proposée en 1935 (voir ci-dessous).

On peut étudier les processus de collision associés à cette interaction. Le problème 4.7 traite de l'annihilation d'un électron et d'un positron vers deux bosons, et se fait à l'aide de la théorie des perturbations ordinaires. On peut aussi démontrer des règles de Feynman particulières à cette théorie. Elles sont un amalgame des règles de Feynman de l'électrodynamique quantiques (section 4.C.2) et de celles valables pour un champ scalaire qui interagit avec lui-même (section 2.C.4). On doit cependant ajouter la possibilité d'un vertex de Yukawa :



$$(4.106)$$

Dans ce vertex, le boson associé au champ scalaire est représenté par une ligne brisée, alors qu'il était représenté par une ligne pleine à la section 2.C.4.

4.E.2 Champ scalaire chargé

Rappelons qu'un champ scalaire ordinaire, obéissant à l'équation de Klein-Gordon, est décrit par le lagrangien suivant :

$$\mathcal{L} = \frac{1}{2} \partial_\mu \phi \partial^\mu \phi - \frac{1}{2} m^2 \phi^2 \quad (4.107)$$

Un tel champ, cependant, ne peut pas être couplé au champ électromagnétique. Pour ce faire, il doit représenter une particule chargée et donc doit nécessairement être complexe (par opposition à réel). En effet, il doit pouvoir admettre une transformation de jauge, qui prend alors la forme

$$\phi \rightarrow \phi' = e^{ie\xi} \phi \quad (4.108)$$

Un champ scalaire complexe peut être vu comme formé de deux champs scalaires réels ϕ_1 et ϕ_2 (l'un pour sa partie réelle, l'autre pour sa partie imaginaire), de sorte que

Prob. 4.2

$$\phi = \frac{1}{\sqrt{2}} (\phi_1 + i\phi_2) \quad (4.109)$$

(le facteur de $\sqrt{2}$ est conventionnel). Le lagrangien du champ scalaire complexe est alors

$$\mathcal{L} = \sum_{i=1,2} \frac{1}{2} \partial_\mu \phi_i \partial^\mu \phi_i - \frac{1}{2} m^2 \phi_i^2 = \partial_\mu \phi^* \partial^\mu \phi - m^2 |\phi|^2 \quad (4.110)$$

Ce champ complexe peut se coupler au champ électromagnétique par couplage minimal, soit en remplaçant $\partial_\mu \phi$ par $\mathcal{D}_\mu \phi$. Le lagrangien complet est alors

$$\mathcal{L} = (\mathcal{D}_\mu \phi)^* \mathcal{D}^\mu \phi - m^2 |\phi|^2 = (\partial_\mu + ieA_\mu) \phi^* (\partial^\mu - ieA^\mu) \phi - m^2 |\phi|^2 \quad (4.111)$$

Notez que le conjugué complexe s'applique à la dérivée covariante, sinon le lagrangien ne serait pas réel. Les termes d'interaction entre le champ électromagnétique et le champ scalaire sont donc

$$\mathcal{L}_I = eA_\mu (i\phi^* \partial^\mu \phi - i\phi \partial^\mu \phi^*) + e^2 |\phi|^2 A_\mu A^\mu \quad (4.112)$$

Le premier de ces termes est cubique (linéaire en A_μ et quadratique en ϕ) et multiplié par e (la charge associée au boson). Le deuxième terme est quartique (quadratique en A_μ et quadratique en ϕ) et multiplié par e^2 . En théorie des perturbations, ceux deux termes sont associés aux vertex suivants :

(i)

$-ie(p_\mu + p'_\mu)$

(ii)

$2ie^2 g_{\mu\nu}$

$$(4.113)$$

Remarques :

- ◆ Le champ scalaire chargé est représenté par une ligne brisée, pour le distinguer d'un fermion, mais ceci est affaire de convention. L'important est que cette ligne comporte une flèche, pour distinguer le boson de son antiparticule.

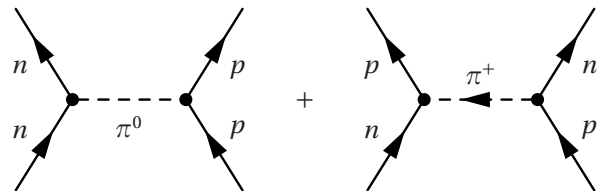
- ◆ Le deuxième vertex porte un facteur e^2 , et donc chaque vertex de ce type compte pour deux ordres de la théorie des perturbations.
- ◆ Dans le premier vertex, les dérivées de ϕ dans le lagrangien correspondant se traduisent par un facteur $p_\mu + p'_\mu$.

On désigne sous le nom d'*électrodynamique scalaire* la théorie de l'interaction d'un champ scalaire chargé avec le champ électromagnétique.

4.E.3 Théorie de Yukawa-Stückelberg de l'interaction forte

Le physicien japonais Hideki YUKAWA et le physicien suisse Ernst STÜCKELBERG⁸ proposèrent indépendamment, en 1935, que la force nucléaire devait sa très courte portée au fait qu'elle était transmise par des particules massives, décrites par un champ scalaire, comme ci-haut. On appela ces particules hypothétiques *mésons*. En supposant que la portée de l'interaction forte est de 1 à 2 fm, on conclut que la masse des mésons doit être située entre 100 et 200 MeV.

Notons qu'il doit y avoir trois types de mésons, de charges électriques $\pm e$ et zéro, qu'on note respectivement π^\pm et π^0 . Le méson neutre π^0 est représenté par un champ scalaire réel et est échangé lors de collisions entre protons ou entre neutrons, comme représenté sur le premier diagramme de l'éq. (4.114). Le méson chargé π^\pm est représenté par un champ scalaire complexe et est échangé entre protons et neutrons, comme représenté dans le deuxième diagramme de l'éq. (4.114) :



$$(4.114)$$

Découverte des mésons La découverte, en 1936, parmi les rayons cosmiques, de particules de masse ~ 105 MeV, fit penser que la particule hypothétique de Yukawa était enfin révélée. On se rendit compte très rapidement que cette particule n'avait pas les propriétés voulues. C'était en fait le *muon*, qu'on appelait initialement le méson- μ . Le véritable méson, le méson- π ou *pion*, fut découvert en 1947. Le méson neutre (π^0) a une masse de 135,0 MeV et le méson chargé une masse de 139,6 MeV. En fait, une pléthore de particules analogues au méson fut découverte dans les années qui suivirent. La théorie de Yukawa n'est manifestement pas une théorie fondamentale de l'interaction forte, quoiqu'elle représente bien certaines de ses propriétés. Fondamentalement, les mésons sont formés de quarks, comme les nucléons. Plus précisément, les

8. Stückelberg est la victime la plus flagrante d'un manque de reconnaissance en physique théorique au 20e siècle. Il est l'auteur de trois découvertes majeures qui ont valu à d'autres des prix Nobel : (1) l'hypothèse du boson intermédiaire pour la force nucléaire; (2) un traitement cohérent des divergences en électrodynamique quantique – la théorie de la renormalisation; (3) la notion de groupe de renormalisation (1954). Dans le premier cas, il a formulé l'hypothèse la même année que Yukawa. Dans le deuxième cas, son travail a été refusé par la *Physical Review* et une version a été publiée dans une revue suisse (en français), des années avant les travaux de Tomonaga, Schwinger et Feynman (1947). Enfin, la notion de groupe de renormalisation et son application à la physique statistique ont valu à K. Wilson le prix Nobel en 1982.

mésons sont formés d'un quark lié à un antiquark, et l'échange de mésons entre deux nucléons peut être vu comme la manifestation de l'échange de quarks entre deux nucléons. Nous reviendrons sur ce sujet dans le chapitre portant sur la classification des hadrons, les particules qui interagissent par l'interaction forte.

F

Problèmes

Question 4.1 Analyse dimensionnelle

Quelles sont les dimensions (au sens des unités) du champ de Dirac, du champ électromagnétique A_μ et du champ scalaire ϕ ? Si un champ scalaire est couplé à un champ de Dirac par le lagrangien d'interaction suivant

$$\mathcal{L}_I = g \phi \psi \bar{\psi} \quad (4.115)$$

quelle est la dimension de la constante de couplage g ? Partez du fait que l'action est sans unité dans le système naturel.

Problème 4.1 Équation de Pauli

Le but de cet exercice est de démontrer que la limite non relativiste de l'équation de Dirac en présence d'un champ électromagnétique est l'équation de Pauli :

$$i\dot{\Psi} = \frac{1}{2m}(\mathbf{P} - e\mathbf{A})^2\Psi - \frac{e}{2m}\mathbf{B} \cdot \boldsymbol{\sigma}\Psi + e\Phi\Psi \quad (4.116)$$

où Ψ est un spineur à deux composantes, Φ est le potentiel électrique et \mathbf{A} le potentiel vecteur. Dans la représentation de Dirac, les deux premières composantes du spineur de Dirac ψ sont plus importantes que les deux dernières, pour les solutions à énergie positive. Supposons en outre que l'énergie de la solution est très proche de m (limite non relativiste). Il est alors sage de définir les spineurs à deux composantes « lents » Ψ et X ainsi :

$$\psi = e^{-imt} \begin{pmatrix} \Psi \\ X \end{pmatrix} \quad (4.117)$$

et de supposer que $X \ll \Psi$.

[A] Montrez que l'équation de Dirac en présence d'un champ électromagnétique quelconque $A_\mu = (\Phi, -\mathbf{A})$ s'exprime ainsi en fonction de Ψ et X :

$$\begin{aligned} i\dot{\Psi} &= \boldsymbol{\pi} \cdot \boldsymbol{\sigma} X + e\Phi\Psi \\ i\dot{X} &= \boldsymbol{\pi} \cdot \boldsymbol{\sigma}\Psi + e\Phi X - 2mX \end{aligned} \quad (4.118)$$

où on a défini la quantité de mouvement $\boldsymbol{\pi} = \mathbf{p} - e\mathbf{A}$.

B Dans la limite non relativiste, supposez que mX est le terme le plus important qui implique X dans la deuxième équation et, par élimination, démontrez l'équation de Pauli pour Ψ . Accessoirement, vous devez démontrer que

$$(\boldsymbol{\pi} \cdot \boldsymbol{\sigma})^2 = \pi^2 - e \mathbf{B} \cdot \boldsymbol{\sigma} \quad (4.119)$$

en vous rappelant que \mathbf{P} est un opérateur différentiel qui agit aussi sur les composantes de \mathbf{A} , lorsque π est élevé au carré. Vous devrez utiliser la formule du produit des matrices de Pauli : $\sigma_a \sigma_b = \delta_{ab} + i \varepsilon_{abc} \sigma_c$ et la relation entre \mathbf{A} et \mathbf{B} .

Problème 4.2 Champ scalaire complexe

En exprimant un champ scalaire complexe en fonction de ses composantes réelles ϕ_1 et ϕ_2 comme à l'éq. (4.109), développez cet champ en fonction des opérateurs de création et d'annihilation de $\phi_{1,2}$ et définissez des opérateurs appropriés pour la création et l'annihilation de particules et d'antiparticules, tels que les deux ont des charges opposées. Basez-vous sur l'expression suivante du courant électromagnétique du champ scalaire complexe :

$$J^\mu = 2 \text{Im}(\phi^* \partial^\mu \phi)$$

Problème 4.3 Quantités physiques et opérateurs de nombre

A La densité d'impulsion du champ électromagnétique est proportionnelle au vecteur de Poynting :

$$\boldsymbol{\pi} = \mathbf{E} \wedge \mathbf{B} \quad (4.120)$$

Montrez qu'on peut exprimer l'impulsion totale du champ de la manière suivante :

$$\mathbf{P} = \sum_{j,\mathbf{k}} \mathbf{k} a_{j\mathbf{k}}^\dagger a_{j\mathbf{k}} \quad (4.121)$$

B Faites de même pour l'impulsion totale des électrons. À partir de l'expression de l'impulsion du champ de Schrödinger, montrez que

$$\mathbf{P} = \sum_{\mathbf{k}} \mathbf{k} c_{\mathbf{k}}^\dagger c_{\mathbf{k}} \quad (4.122)$$

C Étant donné l'expression de la densité d'impulsion du champ électromagnétique, il est naturel de poser que le moment cinétique du champ est donné par l'expression suivante :

$$\mathbf{J} = \int d^3r \, \mathbf{r} \wedge (\mathbf{E} \wedge \mathbf{B}) \quad (4.123)$$

Montrez que cette expression est équivalente à l'expression suivante, si on ne tient compte que de la partie transverse du champ électrique :

$$\mathbf{J} = \int d^3r \left\{ \mathbf{E} \wedge \mathbf{A} + \sum_i E_i (\mathbf{r} \wedge \nabla) A_i \right\} \quad (4.124)$$

On associe le dernier terme au moment cinétique orbital du champ, alors que le premier terme est le moment cinétique intrinsèque (spin). Indice : utilisez la notation indicielle et la formule du double produit vectoriel dans cette notation. Une intégration par partie est nécessaire.

D Exprimez le premier terme en fonction des opérateurs de création et d'annihilation. Utilisez à cette fin une base de polarisations circulaires, c'est-à-dire

$$\mathbf{e}_1 = \frac{\mathbf{x} + i\mathbf{y}}{\sqrt{2}} \quad \mathbf{e}_2 = \frac{\mathbf{x} - i\mathbf{y}}{\sqrt{2}} \quad (4.125)$$

où l'axe des z est dans la direction du vecteur d'onde. Interprétez votre résultat : quel est le spin du photon?

Problème 4.4 Diffusion électron-muon

Ce problème vise à compléter le calcul de la section différentielle de la diffusion entre deux fermions de types différents (par ex. électron sur muon, électron sur proton, etc.). Un seul diagramme de Feynman contribue au processus et est illustré dans les notes ; les quadri-impulsions seront numérotées conformément à cette illustration.

A À partir des règles de Feynman, démontrez l'expression suivante pour l'amplitude du processus :

$$\mathcal{M} = \frac{4\pi\alpha}{(p_1 - p_3)^2} g_{\mu\nu} [\bar{u}_{p_3, s_3} \gamma^\mu u_{p_1, s_1}] [\bar{u}_{p_4, s_4} \gamma^\nu u_{p_2, s_2}] \quad (4.126)$$

B Démontrez la relation

$$\sum_{s=1,2} (u_{\mathbf{p}, s})_\alpha (\bar{u}_{\mathbf{p}, s})_\beta = (p_\mu \gamma^\mu + m)_{\alpha\beta} \quad (4.127)$$

C Démontrez que le carré de l'amplitude de diffusion, sommé sur les spins finaux et moyenné sur les spins initiaux des particules, est

$$\langle |\mathcal{M}|^2 \rangle = \frac{8 \cdot (4\pi\alpha)^2}{(p_1 - p_3)^4} [(p_1 \cdot p_2)(p_3 \cdot p_4) + (p_1 \cdot p_4)(p_2 \cdot p_3) - m^2(p_1 \cdot p_3) - M^2(p_2 \cdot p_4) + 2m^2 M^2] \quad (4.128)$$

D Considérez maintenant le cas où la particule de masse M est très massive en comparaison de l'énergie de la particule de masse m , mais que cette dernière peut tout de même être relativiste. Plaçons-nous dans le référentiel de la cible (particule de masse M au repos). Dans ce cas,

l'énergie transférée à la cible est négligeable, comme si elle était fixe. Démontrez que la section différentielle de diffusion a la forme suivante :

$$\frac{d\sigma}{d\Omega} = \frac{\alpha^2 E^2}{4(\mathbf{p}^2)^2 \sin^4 \theta / 2} (1 - \beta^2 \sin^2 \theta / 2) \quad (4.129)$$

où β est la vitesse du projectile. Ceci est la formule de Mott. Vous devez pour cela utiliser le résultat suivant pour la section différentielle de diffusion dans le repère du laboratoire :

$$\frac{d\sigma}{d\Omega} = \frac{1}{(8\pi)^2} \frac{\mathbf{p}_f^2 |\mathcal{M}_{fi}|^2}{M |\mathbf{p}_i| (|\mathbf{p}_f| (E_i + M) - |\mathbf{p}_i| E_f \cos \theta)} \quad (4.130)$$

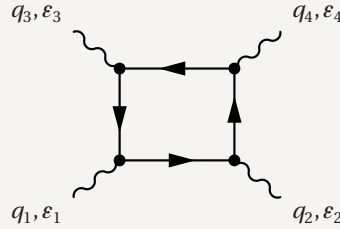
où les indices i et f font référence à l'état initial et l'état final du projectile (l'électron) et M est la masse de la cible (le muon).

Problème 4.5 Corrections radiatives à l'effet Compton

Dessinez tous les diagrammes contribuant à la diffusion électron-photon (effet Compton) au quatrième ordre de la théorie des perturbations.

Problème 4.6 Diffusion photon-photon

Le diagramme suivant décrit le processus de diffusion photon-photon, par l'intermédiaire de fermions virtuels.



A Dessinez tous les diagrammes inéquivalents qui, comme celui-ci, contribuent à ce processus au même ordre de la théorie des perturbations. Indice : il y en a six.

B Écrivez l'expression analytique de ce diagramme, en suivant les règles de Feynman.

C L'intégrale associée à la boucle est-elle convergente ?

Problème 4.7 couplage de Yukawa

Nous allons étudier l'interaction des électrons et des positrons, décrits par le champ de Dirac, avec les bosons décrits par le champ scalaire. Il s'agit ici d'une version simplifiée de l'électrodynamique

quantique (QED) où le photon a été remplacé par un boson sans polarisation et sans masse. Le hamiltonien d'interaction $H_{e\phi}$ entre le champ de Dirac et le champ scalaire doit, pour respecter l'invariance de Lorentz, avoir la forme suivante :

$$H_{e\phi} = g \int d^3r \bar{\psi} \psi \phi \quad (4.131)$$

où $\bar{\psi} = \psi^\dagger \gamma^0$.

A Écrivez une expression du hamiltonien d'interaction $H_{e\phi}$ en fonction des opérateurs de création et d'annihilation des électrons ($c_{\mathbf{p},s}$), des positrons ($d_{\mathbf{p},s}$) et des bosons ($a_{\mathbf{q}}$). Laissez les spins (u, \bar{u}, v, \bar{v}) intacts dans l'expression.

Considérons un processus de collision électron-positron au cours duquel l'électron annihile le positron pour donner deux bosons. Soit \mathbf{p}_1 et \mathbf{p}_2 les quantités de mouvement de l'électron et du positron, respectivement. Les quantités de mouvement des deux bosons de l'état final sont notées \mathbf{p}_3 et \mathbf{p}_4 .

B Montrez que l'amplitude de diffusion M_{fi} s'annule au premier ordre de la théorie des perturbations, mais pas au deuxième ordre et que quatre états intermédiaires y contribuent. Représentez graphiquement les termes correspondants, comme sur la figure 4.1.

C Les quatre termes peuvent être groupés deux par deux : les deux derniers étant obtenus des deux premiers par un échange des deux bosons dans l'état final, de manière similaire à l'exemple discuté dans les notes (chap. 5). Montrez que les deux premiers termes peuvent être combinés de manière à obtenir

$$\frac{g^2}{2\mathcal{V}\sqrt{\omega_3\omega_4}} \bar{v}_{\mathbf{p}_2,s_2} \frac{q \cdot \gamma + m}{q^2 - m^2} u_{\mathbf{p}_1,s_1} \quad (4.132)$$

où $q = p_1 - p_3$. Ici, p_i désigne le quadrivecteur énergie-impulsion de la particule i . L'expression $q \cdot \gamma$ signifie $q^0 \gamma^0 - \sum_{i=1}^3 q^i \gamma^i$.

Note : vous devrez utiliser les formules suivantes pour les sommes sur les spins :

$$\begin{aligned} \sum_{s=1,2} v_{\mathbf{k},s}^\alpha \bar{v}_{\mathbf{k},s}^\beta &= \frac{1}{2\varepsilon_{\mathbf{k}}} (k \cdot \gamma - m)^{\alpha\beta} & (k^0 = \varepsilon_{\mathbf{k}}) \\ \sum_{s=1,2} u_{\mathbf{k},s}^\alpha \bar{u}_{\mathbf{k},s}^\beta &= \frac{1}{2\varepsilon_{\mathbf{k}}} (k \cdot \gamma + m)^{\alpha\beta} & (k^0 = \varepsilon_{\mathbf{k}}) \end{aligned}$$

Les deux derniers termes, quant à eux, sont obtenus en remplaçant q par $\tilde{q} = p_1 - p_4$. L'amplitude totale est donc donnée par l'expression

$$\mathcal{M}_{fi} = \frac{g^2}{2\mathcal{V}\sqrt{\omega_3\omega_4}} \bar{v}_{\mathbf{p}_2,s_2} A u_{\mathbf{p}_1,s_1} \quad \text{où} \quad A = \frac{q \cdot \gamma + m}{q^2 - m^2} + \frac{\tilde{q} \cdot \gamma + m}{\tilde{q}^2 - m^2} \quad (4.133)$$

Pour calculer la section efficace, nous devons mettre cette amplitude au carré. Si l'électron et le positron ne sont pas préparés dans des états précis de spin, il faut sommer le résultat sur les différentes valeurs possibles du spin, c'est-à-dire sur s_1 et s_2 .

D Montrez que

$$\sum_{s_1, s_2} |\mathcal{M}_{fi}|^2 = \frac{g^4}{16\gamma^2 \omega_3 \omega_4 \varepsilon_1 \varepsilon_2} \text{tr}((p_2 \cdot \gamma - m)A(p_1 \cdot \gamma + m)A) \quad (4.134)$$

E Placez-vous dans le référentiel du centre de masse de l'électron et du positron. Soit E l'énergie de chacune des particules incidentes et sortantes et posons $\mathbf{p}_1 = p\mathbf{z}$, $\mathbf{p}_3 = E(\cos\theta\mathbf{z} + \sin\theta\mathbf{x})$. À l'aide de Mathematica, complétez le calcul et démontrez que la section différentielle de diffusion est

$$\frac{d\sigma}{d\Omega} = \frac{g^4}{2048\pi^2 m^2} \frac{x}{\sqrt{1+x^2}} \frac{20+x^2-12\cos 2\theta-x^2\cos 4\theta}{(1+x^2\sin^2\theta)^2} \quad x = \frac{|\mathbf{p}_1|}{m} \quad (4.135)$$

Indice : définissez les matrices de Dirac dans Mathematica, ainsi que des quadrivecteurs et un tenseur métrique. Vous aurez besoin d'une formule démontrée dans les notes pour la section différentielle dans le repère du centre de masse, formule qui demande d'utiliser la normalisation relativiste des états. Vous aurez donc besoin de multiplier l'amplitude au carré par le facteur approprié pour passer à cette normalisation relativiste.

F Illustrez la dépendance angulaire de la section différentielle (c.-à-d. un graphique en fonction de θ de $\theta = 0$ à $\theta = \pi$ pour $p/m = 0, 1, 1$ et 10 . Que remarquez-vous?

Problème 4.8 couplage de Yukawa (avec règles de Feynman)

Refaites le problème 4.7, cette fois à l'aide des règles de Feynman, qui sont formulées dans la normalisation relativiste des états. Autrement dit : trouver l'amplitude de diffusion \mathcal{M} pour l'annihilation d'un électron et d'un positron en deux bosons de spin zéro mais de masse nulle.

Question 4.2 Théorie de Yukawa-Stückelberg

Écrivez une expression possible pour le lagrangien d'interaction \mathcal{L}_I entre les champs de Dirac représentant les neutrons (ψ_n), les protons (ψ_p) et les champs scalaires représentant le pion neutre (ϕ_0) et le pion chargé (ϕ). N'oubliez pas que ϕ_0 est réel, alors que ϕ est complexe. Cette expression doit mener aux diagrammes de l'éq. (4.114).

CHAPITRE 5

SYMÉTRIES

La symétrie est une caractéristique centrale des théories des particules élémentaires. On peut affirmer que ces théories ont évolué au cours du XXe siècle en donnant à la symétrie un rôle de plus en plus important. En particulier, les théories des interactions fondamentales (forte, électromagnétique et faible) sont basées sur le concept de symétrie locale, ou symétrie de jauge. Également importante est la manière dont certaines symétries, présentent à un niveau fondamental, sont dissimulées dans les faits par un phénomène appelé « brisure spontanée de la symétrie ». Même imparfaites, les symétries constituent un aspect extrêmement important de notre compréhension de l'Univers.

Théorème de Wigner Une opération de symétrie est une transformation affectant une ou plusieurs quantités physiques, ou observables (en mécanique quantique). C'est Eugene WIGNER qui, le premier, en 1931, énonça les conséquences générales des symétries en mécanique quantique. Soit une opération de symétrie générale, dont l'effet sur les états quantiques est représenté par un opérateur $S : |\psi\rangle$ représentant le nouvel état quantique, suite à l'application de la transformation (par exemple, une rotation, une translation, ou une réflexion). L'exigence de symétrie revient à poser que, pour toute paire d'états $|\psi_1\rangle$ et $|\psi_2\rangle$, on ait

$$|\langle S\psi_2 | S\psi_1 \rangle|^2 = |\langle \psi_2 | \psi_1 \rangle|^2 \quad (5.1)$$

Le théorème de Wigner¹ stipule que S est un opérateur soit unitaire ou antiunitaire :

$$\begin{aligned} (i) : \langle S\psi_2 | S\psi_1 \rangle &= \langle \psi_2 | \psi_1 \rangle \implies S^\dagger S = 1 && \text{(unitaire)} \\ (ii) : \langle S\psi_2 | S\psi_1 \rangle &= \langle \psi_1 | \psi_2 \rangle = \langle \psi_2 | \psi_1 \rangle^* = 1 && \text{(antiunitaire)} \end{aligned}$$

Par définition, une transformation $f(A)$ agissant sur un espace vectoriel est *antilinéaire* si $f(\lambda A + \eta B) = \lambda^* f(A) + \eta^* f(B)$. En particulier, une transformation est *antiunitaire* si $\langle f(A) | f(B) \rangle = \langle B | A \rangle = \langle A | B \rangle^*$. Le produit de deux transformations antilinéaires est une transformation linéaire. On peut tout de même représenter une transformation antilinéaire par une matrice, mais ce choix de matrice non seulement dépend de la base choisie, mais ne se transforme pas comme une matrice lorsqu'on procède à un changement de base.

1. E.P. Wigner, *Gruppentheorie und ihre Anwendung auf die Quantenmechanik des Atomspektren*, Braunschweig 1931. Pour une démonstration plus complète, voir S. Weinberg, *the quantum theory of fields*, Cambridge (1995), annexe 2.A.

A C, P et T

Parmi les symétries de la nature, trois jouent un rôle particulier, du fait de leur simplicité :

1. La symétrie miroir, par laquelle on effectue une réflexion de l'espace, comme si l'Univers était examiné dans un miroir. En trois dimensions, cette réflexion est équivalente à une inversion des coordonnées $\mathbf{r} \mapsto -\mathbf{r}$; elle porte aussi le nom de *transformation de parité* et est notée par la lettre P .
2. L'échange entre matière et antimatière, aussi appelé *conjugaison de charge* et noté C .
3. L'*inversion du temps*, ou *inversion du mouvement*, qui est obtenue en changeant le signe du temps t dans les équations microscopiques. Cette opération est notée T .

Ces trois opérations de symétrie discrètes ont ceci de particulier que leur carré est égal à l'unité : $P^2 = C^2 = T^2 = 1$. On peut aussi démontrer que le produit CPT des trois opérations est une symétrie exacte de toute théorie raisonnable, même si les symétries individuelles P , C et T peuvent être brisées par certaines interactions fondamentales (en particulier les interactions faibles). Nous allons commencer ce chapitre en étudiant ces trois symétries à tour de rôle.

5.A.1 La parité, ou inversion de l'espace (P)

Définition En général, la transformation de parité est définie par la réflexion de l'espace par rapport à un plan donné, par exemple le plan $x = 0$. Son effet sur les coordonnées dans l'espace-temps serait donc

$$(t, x, y, z) \mapsto (t, -x, y, z) \quad (\text{réflexion}) \quad (5.2)$$

Comme on suppose généralement que les lois fondamentales de la Nature soit invariantes par rotation dans l'espace, on peut ajouter à cette réflexion, en trois dimensions, une rotation de π par rapport à l'axe des x , ce qui fait qu'une transformation équivalente est l'inversion de l'espace :

$$(t, x, y, z) \mapsto (t, -x, -y, -z) \quad (\text{inversion}) \quad (5.3)$$

Cette équivalence n'est cependant pas valable en deux dimensions, ou plus généralement dans un nombre pair de dimensions spatiales.

Rotations et pseudo-rotations Une rotation est définie comme une transformation préservant la norme des vecteurs *et* qui peut être obtenue de manière continue de la transformation identité. On montre facilement qu'une rotation est caractérisée par une matrice orthogonale R , donc une matrice telle que $\tilde{R}R = \mathbb{I}$. Cependant, toutes les matrices orthogonales ne sont pas des rotations : la matrice $-\mathbb{I}$, qui effectue une inversion de l'espace, est orthogonale, et toute matrice de la forme $-R$, où R est une matrice de rotation, est aussi orthogonale. Plus précisément, une matrice orthogonale O satisfait nécessairement à la propriété $\det(\tilde{O}O) = (\det O)^2 = 1$, d'où $\det O = \pm 1$. Les matrices R dont le déterminant est 1 peuvent être obtenues de manière continue de la matrice identité et correspondent aux vraies rotations. Les autres

peuvent toujours être mises sous la forme $-R$, où R est une vraie rotation, et constituent ce qu'on appelle des *pseudo-rotations*.

Scalaire, pseudo-scalaire, vecteurs polaires et axiaux Généralement, une quantité invariante par rotation est qualifiée de *scalaire*, alors qu'une quantité qui se transforme comme les coordonnées (x, y, z) lors d'une rotation est qualifiée de *vecteur*. Cette définition doit être nuancée de la manière suivante : un vrai scalaire est invariant suite à une rotation ou à une pseudo-rotation. Par contre, un *pseudo-scalaire* est une quantité qui, tout en étant invariante suite à une vraie rotation, change de signe suite à une pseudo-rotation, c'est-à-dire une inversion de l'espace.

De même, un *vecteur polaire* se transforme comme les coordonnées suite à une rotation ou une pseudo-rotation ; en particulier, il change de signe suite à une inversion de l'espace. Par contre, un *vecteur axial* se transforme comme un vecteur suite à une rotation, mais ne change pas de signe suite à une inversion. Parmi les vecteurs polaires, on note la position \mathbf{r} et ses dérivées par rapport au temps (vitesse, accélération), l'impulsion \mathbf{p} , le champ électrique \mathbf{E} . Les vecteurs axiaux, quant à eux, impliquent généralement un produit vectoriel : le moment cinétique \mathbf{J} (orbital ou de spin), le champ magnétique \mathbf{B} . Le produit scalaire d'un vecteur polaire par un vecteur axial, par exemple $\mathbf{E} \cdot \mathbf{B}$, est un pseudo-scalaire.

Opération de parité en mécanique quantique Selon les règles générales de la mécanique quantique, cette transformation de symétrie doit correspondre à un opérateur unitaire Π agissant dans l'espace des phases, tel que $\Pi^2 = 1$ et donc tel que $\Pi^\dagger = \Pi$. Suite à une telle transformation, le hamiltonien H est généralement modifié comme suit :

$$H' = \Pi H \Pi \quad (5.4)$$

Si le système décrit par le hamiltonien H est invariant par inversion de l'espace, alors $H' = H$ et l'opérateur Π commute avec le hamiltonien : $[H, \Pi] = 0$. Les états propres du hamiltonien sont alors des états propres de Π . Comme $\Pi^2 = 1$, les valeurs propres de Π sont ± 1 .

Dans le cadre de la mécanique quantique non relativiste à une particule, un système invariant par inversion est souvent décrit par un potentiel $V(\mathbf{r})$ tel que $V(-\mathbf{r}) = V(\mathbf{r})$. En particulier, en une dimension d'espace, un potentiel symétrique $V(-x) = V(x)$ mène à des solutions à l'équation de Schrödinger qui sont soit paires ($\Pi|\psi\rangle = |\psi\rangle$) ou impaires ($\Pi|\psi\rangle = -|\psi\rangle$).

L'opérateur du moment cinétique \mathbf{L} étant un vecteur axial, on a la relation opératorielle $\Pi \mathbf{L} \Pi = \mathbf{L}$, ou encore $[\Pi, \mathbf{L}] = 0$. Ceci signifie que les états propres d'une composante de \mathbf{L} auront aussi une parité bien définie (ce seront des états propres de Π). On montre sans peine, d'après l'expression des harmoniques sphériques, qu'une fonction d'onde appartenant à un multiplet de nombre quantique orbital l a une parité $(-1)^l$.

Parité du champ de Dirac Quel est l'effet de l'inversion de l'espace sur un spineur de Dirac ψ ? Rappelons que, dans la représentation chirale, le spineur de Dirac à quatre composantes est formé des spineurs droit et gauche à deux composantes de la manière sui-

vante :

$$\psi = \begin{pmatrix} \chi_L \\ \chi_R \end{pmatrix} \quad (\text{rep. chirale}) \quad (5.5)$$

(dorénavant la notation $\chi_{L,R}$ est utilisée pour les spineurs de Weyl). On est donc en droit de supposer que l'opération d'inversion va simplement échanger les deux spineurs de Weyl, ce qui est obtenu d'une multiplication par la matrice γ^0 (la représentation chirale des matrices de Dirac est la plus appropriée pour cet exercice, mais le résultat est indépendant de la représentation). On définit donc l'opération d'inversion sur un spineur de Dirac comme

$$\psi(\mathbf{r}, t) \rightarrow \psi'(\mathbf{r}, t) = \eta \gamma^0 \psi(-\mathbf{r}, t) \quad (5.6)$$

où $\eta = \pm 1$ est la *parité intrinsèque* du champ de Dirac considéré. Un état de fermion ou d'antifermion avec vecteur d'onde \mathbf{k} non nul n'est pas un état propre de la parité, puisque \mathbf{k} serait transformé en $-\mathbf{k}$ suite à une inversion de l'espace. Par contre, si on considère les états représentant des particules au repos, alors on voit clairement, à partir de la représentation de Dirac de la matrice γ^0 et les équations (3.101) que les antifermions ont une parité opposée à celle des fermions.

Composantes droite et gauche d'un spineur de Dirac Dans le but d'extraire ces composantes dans le cadre de spineurs à quatre composantes, on introduit la matrice suivante :

$$\gamma^5 = i\gamma^0\gamma^1\gamma^2\gamma^3 \quad (5.7)$$

La forme explicite de cette matrice est la suivante, dans les deux représentations utilisées des matrices de Dirac :

$$\gamma^5 = \begin{pmatrix} -1 & 0 \\ 0 & 1 \end{pmatrix} \quad (\text{rep. chirale}) \quad \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \quad (\text{rep. de Dirac}) \quad (5.8)$$

On montre facilement les propriétés suivantes de γ^5 :

$$\{\gamma^5, \gamma^\mu\} = 0 \quad (\gamma^5)^2 = 1 \quad \gamma^0\gamma^5\gamma^0 = (\gamma^5)^\dagger \quad (5.9)$$

On définit ensuite les composantes droite et gauche du spineur de Dirac comme

$$\boxed{\psi_R = \frac{1}{2}(1 + \gamma^5)\psi \quad \psi_L = \frac{1}{2}(1 - \gamma^5)\psi} \quad (5.10)$$

En effet, dans la représentation chirale,

$$\begin{aligned} \psi_R &= \frac{1}{2}(1 + \gamma^5)\psi = \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix} \psi = \begin{pmatrix} 0 \\ \chi_R \end{pmatrix} \\ \psi_L &= \frac{1}{2}(1 - \gamma^5)\psi = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} \psi = \begin{pmatrix} \chi_L \\ 0 \end{pmatrix} \end{aligned} \quad (5.11)$$

Notons que les spineurs droit et gauche à 4 composantes (ψ_R et ψ_L) contiennent la même information que les spineurs de Weyl à deux composantes correspondants (χ_R et χ_L). Il est toutefois utile

de disposer de cette représentation à 4 composantes pour les spineurs chiraux. Notons aussi que la définition (5.10) est valable dans toutes les représentations des matrices de Dirac ; c'est la forme matricielle précise de γ^5 qui change d'une représentation à l'autre.

Les spineurs conjugués droit et gauche sont obtenus en changeant le signe de γ^5 :

$$\boxed{\bar{\psi}_R = \bar{\psi} \frac{1}{2}(1 - \gamma^5) \quad \bar{\psi}_L = \bar{\psi} \frac{1}{2}(1 + \gamma^5)} \quad (5.12)$$

En effet, toujours dans la représentation chirale,

$$\begin{aligned} \bar{\psi}_R &= \bar{\psi} \frac{1}{2}(1 - \gamma^5) = \psi^\dagger \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} = (\chi_R^\dagger, 0) \\ \bar{\psi}_L &= \bar{\psi} \frac{1}{2}(1 + \gamma^5) = \psi^\dagger \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix} = (0, \chi_L^\dagger) \end{aligned} \quad (5.13)$$

En fonction de ces spineurs chiraux, on vérifie facilement que l'action de Dirac a la forme suivante :

$$\mathcal{L}_D = i\bar{\psi}_L \gamma^\mu \partial_\mu \psi_L + i\bar{\psi}_R \gamma^\mu \partial_\mu \psi_R - m(\bar{\psi}_L \psi_R + \bar{\psi}_R \psi_L) \quad (5.14)$$

On constate que le terme de masse couple les spineurs gauches et droits, alors que le reste du lagrangien se sépare bien en deux termes, l'un gauche et l'autre droit.

Dans le cas sans masse, les spineurs droit et gauche décrivent des particules d'hélicité bien définie, comme nous l'avons entrevu à la section 3.A.5. Dans ce cas, les spineurs droit et gauche peuvent être considérés séparément, c'est-à-dire qu'on peut définir une théorie ne comportant que la composante gauche ou que la composante droite, car les deux ne sont pas couplées par le terme de masse et donc sont indépendantes. Dans le cas massif, les composantes chirales ne décrivent plus des particules d'hélicité bien définie, mais demeurent utiles dans la description d'interactions qui brisent la parité, comme nous le verrons au chapitre suivant.

Étant donnés deux spineurs de Dirac ψ_1 et ψ_2 , la matrice γ^5 nous permet de définir un pseudo-scalaire et un pseudo-vecteur :

$$\bar{\psi}_1 \gamma^5 \psi_2 \quad \text{et} \quad \bar{\psi}_1 \gamma^\mu \gamma^5 \psi_2 \quad (5.15)$$

Ces deux quantités se transforment respectivement comme un scalaire et un vecteur lors des transformations de Lorentz qui n'impliquent pas d'inversion de l'espace. On vérifie, à l'aide de la forme donnée ci-haut de la transformation de parité sur un spineur de Dirac, que la première quantité ($\bar{\psi}_1 \gamma^5 \psi_2$) change de signe lors d'une réflexion de l'espace : c'est un *pseudo-scalaire*. De même, la partie spatiale de la deuxième quantité ($\bar{\psi}_1 \gamma^\mu \gamma^5 \psi_2$) est un vecteur *axial*.

Parité du champ électromagnétique Le champ électromagnétique A_μ se transforme comme suit lors d'une inversion de l'espace :

$$A'_0(\mathbf{r}, t) = A_0(-\mathbf{r}, t) \quad \mathbf{A}'(\mathbf{r}, t) = -\mathbf{A}(-\mathbf{r}, t) \quad (5.16)$$

Autrement dit, le potentiel électrique A_0 est un scalaire et le potentiel vecteur est un vecteur polaire, en raison de la relation de ces quantités avec les champs électrique (vecteur polaire) et magnétique (vecteur axial) :

$$\mathbf{E} = -\nabla A_0 - \frac{\partial \mathbf{A}}{\partial t} \quad \mathbf{B} = \nabla \wedge \mathbf{A} \quad (5.17)$$

Étant donné un quadrivecteur A_μ , on désigne par \tilde{A}_μ le quadrivecteur obtenu en inversant les composantes spatiales seulement :

$$\tilde{A}_0 = A_0 \quad \tilde{A}_i = -A_i \quad (i = 1, 2, 3) \quad (5.18)$$

L'effet de l'inversion de l'espace sur le champ électromagnétique peut donc s'écrire

$$A_\mu(\mathbf{r}, t) \rightarrow A'_\mu(\mathbf{r}, t) = \tilde{A}_\mu(-\mathbf{r}, t) \quad (5.19)$$

Notons qu'un état à un photon de vecteur d'onde déterminé n'est pas invariant par inversion de l'espace, puisque $\mathbf{k} \mapsto -\mathbf{k}$. Ces états n'ont donc pas une parité bien déterminée. Par contre, une onde sphérique, développée en harmoniques sphériques, est un état propre de la parité. On montre, par exemple, qu'un photon émis radialement par rayonnement dipolaire a une parité négative.

Invariance de l'équation de Dirac L'équation de Dirac en présence d'un champ électromagnétique prend la forme suivante :

$$i\gamma^\mu \partial_\mu \psi - e\gamma^\mu A_\mu \psi - m\psi = 0 \quad (5.20)$$

Vérifions que cette équation est inchangée lorsqu'on remplace les champs $\psi(\mathbf{r}, t)$ et $A_\mu(\mathbf{r}, t)$ par les champs transformés par parité selon les équations (5.6) et (5.19). Nous allons supposer que l'équation est encore valable pour les champs transformés et vérifier à la fin que cette hypothèse est compatible avec l'équation de Dirac originale :

$$\begin{aligned} i\gamma^\mu \partial_\mu \psi'(\mathbf{r}, t) - e\gamma^\mu A'_\mu(\mathbf{r}, t) \psi'(\mathbf{r}, t) - m\psi'(\mathbf{r}, t) &= 0 \\ &= i\eta\gamma^\mu \gamma^0 \partial_\mu \psi(-\mathbf{r}, t) - e\eta\gamma^\mu \tilde{A}_\mu(-\mathbf{r}, t) \gamma^0 \psi(-\mathbf{r}, t) - m\eta\gamma^0 \psi(-\mathbf{r}, t) \\ &= i\eta\gamma^\mu \gamma^0 \tilde{\partial}_\mu \psi(\mathbf{r}, t) - e\eta\gamma^\mu \gamma^0 \tilde{A}_\mu(\mathbf{r}, t) \psi(\mathbf{r}, t) - m\eta\gamma^0 \psi(\mathbf{r}, t) \\ &= i\gamma^0 \gamma^\mu \gamma^0 \tilde{\partial}_\mu \psi(\mathbf{r}, t) - e\gamma^0 \gamma^\mu \gamma^0 \tilde{A}_\mu(\mathbf{r}, t) \psi(\mathbf{r}, t) - m\psi(\mathbf{r}, t) \end{aligned} \quad (5.21)$$

À la deuxième ligne nous avons procédé au changement $\mathbf{r} \rightarrow -\mathbf{r}$, ce qui entraîne un changement correspondant des dérivées $\partial_\mu \rightarrow \tilde{\partial}_\mu$. À la troisième ligne nous avons prémultiplié par γ^0 et divisé par η . Étant données les relations d'anticommutation des matrices de Dirac, on constate facilement que

$$\gamma^0 \gamma^\mu \gamma^0 = \tilde{\gamma}^\mu \quad (5.22)$$

et donc on obtient l'équation suivante :

$$i\tilde{\gamma}^\mu \tilde{\partial}_\mu \psi(\mathbf{r}, t) - e\tilde{\gamma}^\mu \tilde{A}_\mu(\mathbf{r}, t) \psi(\mathbf{r}, t) - m\psi(\mathbf{r}, t) = 0 \quad (5.23)$$

Or, évidemment, pour tout produit de quadrivecteurs, $\tilde{A}_\mu \tilde{B}^\mu = A_\mu B^\mu$ et donc l'équation ci-dessus devient identique à l'équation de Dirac originale. Donc l'équation de Dirac demeure inchangée lors d'une transformation de parité affectant à la fois le champ électronique ψ et le champ électromagnétique A_μ .

5.A.2 La conjugaison de charge (C)

Définition L'opération de conjugaison de charge est définie intuitivement comme l'échange des particules et des antiparticules. Nous allons définir son effet sur un spineur de Dirac et sur le champ électromagnétique, comme dans la dernière partie de la section précédente portant sur la parité, et ensuite vérifier que cette transformation laisse l'équation de Dirac inchangée.

$$\psi \rightarrow \psi^c = i\eta_c \gamma^2 \psi^* = \eta_c \mathcal{C} \bar{\psi}^T \quad A_\mu \rightarrow A_\mu^c = -A_\mu \quad (5.24)$$

où la matrice \mathcal{C} est $i\gamma^2\gamma^0$ et $\eta_c = \pm 1$, selon la particule impliquée. Comme cette transformation n'affecte aucunement l'espace et le temps, nous n'écrivons pas explicitement les arguments (\mathbf{r}, t) des champs impliqués.

Effet sur l'équation de Dirac Appliquons une conjugaison complexe à l'équation de Dirac en présence d'un champ électromagnétique :

$$(-i\partial_\mu - eA_\mu)\gamma^{\mu*}\psi^* - m\psi^* = 0 \quad (5.25)$$

Multiplions ensuite par γ^2 , en insérant $-(\gamma^2)^2 = 1$ devant ψ^* dans le premier terme :

$$(i\partial_\mu + eA_\mu)\gamma^2\gamma^{\mu*}\gamma^2\psi^* - m\gamma^2\psi^* = 0 \quad (5.26)$$

Il reste à calculer $\gamma^2\gamma^{\mu*}\gamma^2$. Signalons avant tout que $\gamma^{2*} = -\gamma^2$. Donc $\gamma^2\gamma^{\mu*}\gamma^2 = (\gamma^2\gamma^\mu\gamma^2)^*$. Si $\mu \neq 2$, l'anticommutation donne $\gamma^2\gamma^\mu\gamma^2 = -\gamma^\mu\gamma^2\gamma^2 = \gamma^\mu$, et ensuite $\gamma^{\mu*} = \gamma^\mu$. Si, au contraire, $\mu = 2$, alors $(\gamma^2\gamma^\mu\gamma^2)^* = -\gamma^{2*} = \gamma^2 = \gamma^\mu$, car γ^2 est purement imaginaire, alors que les autres matrices de Dirac sont réelles. On trouve donc $\gamma^2\gamma^{\mu*}\gamma^2 = \gamma^\mu$ et

$$(i\partial_\mu + eA_\mu)\gamma^\mu(\gamma^2\psi^*) - m\gamma^2\psi^* = 0 \quad (5.27)$$

En multipliant par i on peut exprimer cette équation en fonction de ψ^c , ce qui donne

$$(i\partial_\mu + eA_\mu)\gamma^\mu\psi^c - m\psi^c = 0 \quad (5.28)$$

On retrouve donc l'équation de Dirac pour ψ^c , mais avec une charge opposée ($e \rightarrow -e$). Pour retrouver l'équation de Dirac originale, il faut aussi appliquer la conjugaison de charge au champ électromagnétique, c'est-à-dire remplacer A_μ par $-A_\mu^c$:

$$(i\partial_\mu - eA_\mu^c)\gamma^\mu\psi^c - m\psi^c = 0 \quad (5.29)$$

La transformation définie en (5.24) est donc une symétrie de l'électrodynamique quantique.

Effet sur les spineurs de base Le fait que la conjugaison de charge appliquée sur le champ de Dirac nécessite l'inversion de la charge électrique ($e \rightarrow -e$) si on ne transforme pas convenablement le champ A_μ confirme faiblement l'interprétation de cette transformation comme un échange des particules et des antiparticules. Pour confirmer cette interprétation plus solidement, on doit appliquer la transformation aux spineurs de base $u_{\mathbf{p},s}$ et $v_{\mathbf{p},s}$. On vérifie sans peine que

$$u_{\mathbf{p},1}^c = v_{\mathbf{p},1} \quad u_{\mathbf{p},2}^c = -v_{\mathbf{p},2} \quad v_{\mathbf{p},1}^c = u_{\mathbf{p},1} \quad v_{\mathbf{p},2}^c = -u_{\mathbf{p},2} \quad (5.30)$$

On constate que les spineurs décrivant les électrons ont été échangés avec ceux décrivant les positrons (modulo un signe sans gravité).

Effet dans l'espace de Hilbert La conjugaison de charge a aussi un effet sur les états quantiques. L'opérateur C de conjugaison de charge peut être défini par son action sur les opérateurs de création et d'annihilation, de manière à calquer son effet (5.30) sur les spineurs :

$$C c_{\mathbf{p},s} C^{-1} = (-1)^{s+1} d_{\mathbf{p},s} \quad C d_{\mathbf{p},s} C^{-1} = (-1)^{s+1} c_{\mathbf{p},s} \quad (5.31)$$

L'opérateur C est unitaire et son carré doit être l'identité, puisque la conjugaison de charge doit nous faire revenir à l'état de départ lorsqu'on l'applique deux fois de suite. Donc $C = C^\dagger = C^{-1}$. Notons donc que la deuxième des équations ci-dessous est une conséquence de la première et que les mêmes équations s'appliquent aux opérateurs de création : $C c_{\mathbf{p},s}^\dagger C^{-1} = (-1)^{s+1} d_{\mathbf{p},s}^\dagger$.

En substituant le développement en modes du champ de Dirac, la correspondance entre $C\psi C$ et ψ^c s'exprime comme suit :

$$\begin{aligned} C\psi C &= \frac{1}{\sqrt{\gamma}} \sum_{\mathbf{p}} \sum_{s=1,2} \left(C c_{\mathbf{p},s} C u_{\mathbf{p},s} e^{i\mathbf{p}\cdot\mathbf{r}} + C d_{\mathbf{p},s}^\dagger C v_{\mathbf{p},s} e^{-i\mathbf{p}\cdot\mathbf{r}} \right) \\ &= \frac{1}{\sqrt{\gamma}} \sum_{\mathbf{p}} \sum_{s=1,2} (-1)^{s+1} \left(d_{\mathbf{p},s} u_{\mathbf{p},s} e^{i\mathbf{p}\cdot\mathbf{r}} + c_{\mathbf{p},s}^\dagger v_{\mathbf{p},s} e^{-i\mathbf{p}\cdot\mathbf{r}} \right) \end{aligned} \quad (5.32)$$

En revanche, en appliquant la conjugaison de charge sur le champ ψ directement, en interprétant la conjugaison complexe comme un conjugué hermitien lorsqu'elle s'applique sur des opérateurs, on trouve

$$\begin{aligned} \psi^c &= \frac{1}{\sqrt{\gamma}} \sum_{\mathbf{p}} \sum_{s=1,2} \left(c_{\mathbf{p},s}^\dagger i\gamma^2 u_{\mathbf{p},s}^* e^{-i\mathbf{p}\cdot\mathbf{r}} + d_{\mathbf{p},s} i\gamma^2 v_{\mathbf{p},s}^* e^{i\mathbf{p}\cdot\mathbf{r}} \right) \\ &= \frac{1}{\sqrt{\gamma}} \sum_{\mathbf{p}} \sum_{s=1,2} (-1)^{s+1} \left(c_{\mathbf{p},s}^\dagger v_{\mathbf{p},s} e^{-i\mathbf{p}\cdot\mathbf{r}} + d_{\mathbf{p},s} u_{\mathbf{p},s} e^{i\mathbf{p}\cdot\mathbf{r}} \right) \end{aligned} \quad (5.33)$$

On constate que les deux expressions $C\psi C$ et ψ^c coïncident, ce qui confirme la définition (5.31) de l'action de l'opérateur C sur les opérateurs de création et d'annihilation.

Notons que la transformation $\psi \rightarrow \psi^c = i\gamma^2 \psi^*$ n'est pas vraiment représentable par une matrice, car c'est une transformation *antilinéaire*.

Remarques :

- ♦ dans une représentation des matrices différente de celle de Dirac, la matrice de conjugaison de charge \mathcal{C} n'aura plus la forme (5.24), soit $\mathcal{C} = i\gamma^2\gamma^0$. Par contre, elle respectera toujours la propriété

$$\mathcal{C} \tilde{\gamma}^\mu \mathcal{C}^{-1} = -\gamma^\mu \quad (5.34)$$

- ♦ Une autre façon de voir comment la parité des antiparticules est l'opposée de celle des particules est de noter que les opérations d'inversion de l'espace et de conjugaison de charge sur un spineur de Dirac anticommulent, en raison de l'anticommutation des matrices γ^0 et γ^2 . En fonction d'opérateurs Π et C agissant dans l'espace de Hilbert, ceci se traduit par

$$\Pi C + C \Pi = 0 \implies C \Pi C = -\Pi \quad (5.35)$$

ce qui signifie que l'opération de conjugaison de charge change la parité d'un état propre de Π .

Conjugaison de charge du photon Une particule qui est sa propre antiparticule est en principe un état propre de l'opérateur C de conjugaison de charge. Comme le champ électromagnétique A_μ ne fait que changer de signe lors de la conjugaison de charge, les états à un photon sont donc des états propres de l'opérateur C , avec valeur propre -1 .

Champ scalaire complexe L'opération de conjugaison de charge sur un champ scalaire complexe ϕ est simplement définie comme une conjugaison complexe, multipliée par une phase intrinsèque :

$$\phi^c = \eta_C \phi^* \quad (5.36)$$

Il est plutôt évident que l'équation de Klein-Gordon pour le champ complexe ne sera pas affectée par cette transformation, et que le couplage minimal au champ électromagnétique restera inchangé si le champ électromagnétique subit aussi la transformation $A_\mu \mapsto A_\mu^c = -A_\mu$.

Exemple 5.1 Désintégration du pion neutre

Le pion neutre π^0 est un état lié d'un quark q et de son antiquark \bar{q} (en fait, principalement la superposition de deux états liés, $u\bar{u}$ et $d\bar{d}$, mais ce détail n'est pas important ici). On est en droit d'affirmer que la fonction d'onde du quark q et de l'antiquark \bar{q} est spatialement symétrique, comme c'est le cas généralement pour l'état fondamental d'un système de deux particules. Par contre, le pion possède un spin nul, et est donc dans un état de singulet de spin, qui est antisymétrique lors de l'échange du quark et de l'antiquark. Symboliquement, l'état de spin du pion est donc

$$|\pi^0\rangle = \frac{1}{\sqrt{2}} (|q_1 \bar{q}_1\rangle - |q_1 \bar{q}_1\rangle) \quad (5.37)$$

L'opération de conjugaison de charge va échanger le quark q de spin s avec l'antiquark \bar{q} de même spin, comme à l'éq. (5.31), de sorte que l'effet net sur l'état du pion est :

$$C|\pi^0\rangle = \frac{1}{\sqrt{2}} (|\bar{q}_1 q_1\rangle - |\bar{q}_1 q_1\rangle) \quad (5.38)$$

Comme les états à plusieurs fermions sont antisymétriques lors de l'échange de deux particules, on retrouve l'état original si on met le quark et l'antiquark dans le même ordre qu'au départ :

$$C|\pi^0\rangle = \frac{1}{\sqrt{2}} (|q_1 \bar{q}_1\rangle - |q_1 \bar{q}_1\rangle) = |\pi^0\rangle \quad (5.39)$$

On conclut que le pion a une valeur propre de C égale à $+1$. Si l'état quark-antiquark était de spin s et de moment cinétique orbital l au lieu de $s = 0$ et $l = 0$ comme le π^0 , la valeur propre de C serait plutôt $(-1)^{s+l}$. Le pion neutre ayant donc $C = 1$, les produits de sa désintégration doivent aussi, au total, avoir $C = 1$. Le photon ayant $C = -1$, le pion ne peut donc pas se désintégrer en un seul photon. Bien sûr, le pion ne pourrait pas se désintégrer en un seul photon pour de simples raisons cinématiques, mais l'interdiction du processus par la conservation de C vaut pour des photons virtuels également, de sorte que la désintégration $\pi^0 \rightarrow e\bar{e}$, via un photon virtuel, est également interdite. Le mode dominant de désintégration du π^0 est en deux photons : $\pi^0 \rightarrow 2\gamma$ (98,8%), alors que le rapport d'embranchement pour $\pi^0 \rightarrow e\bar{e}$ est de 6×10^{-8} , signe d'une très faible violation de la symétrie C dans ce processus (vraisemblablement par l'intermédiaire des interactions faibles).

À noter que la parité du π^0 doit être -1 , par l'argument suivant : l'inversion de l'espace se trouve à inverser la coordonnée relative du quark et de son antiquark, ce qui donne un signe $+1$ car la fonction d'onde est

symétrique dans l'état fondamental. Par contre, la parité intrinsèque de l'antiquark étant l'opposée de celle du quark, l'effet net de la transformation de parité sur le π^0 est -1 . Le pion (chargé ou neutre) est donc une particule *pseudo-scalaire*, c'est-à-dire que le champ décrivant cette particule devrait être un pseudo-scalaire et non un champ scalaire. Si le moment orbital de la paire quark-antiquark était l , alors la parité serait $(-1)^{l+1}$. La même combinaison de quark-antiquark avec $l = 1$ représente en fait un méson de spin $J = 1$ formant un *vecteur polaire*, le ρ .

5.A.3 L'inversion du temps (T)

La dernière des trois grandes symétries discrètes est l'inversion du temps, qui revient à changer t en $-t$ dans les équations fondamentales. L'inversion du temps est parfois appelée *inversion du mouvement*, car elle revient à passer le film des événements à l'envers. Les quantités qui changent de signe lors de l'inversion du temps sont celles qui sont habituellement associées à une vitesse, comme l'impulsion \mathbf{p} , le moment cinétique \mathbf{J} et le champ magnétique \mathbf{B} . Comme l'irréversibilité est une caractéristique fondamentale des phénomènes complexes, il semble que l'inversion du cours du temps ne soit pas une symétrie de la nature. Par contre, sauf une très rare exception qui sera discutée au chapitre suivant, les phénomènes microscopiques semblent respecter cette symétrie. C'est vrai en mécanique classique si on excepte les processus dissipatifs, et c'est vrai en mécanique quantique si on assortit l'inversion d'une conjugaison complexe.

Invariance de l'équation de Dirac L'effet de l'inversion du temps sur le champ de Dirac et sur le champ électromagnétique peut être définie de la manière suivante :

$$\psi(\mathbf{r}, t) \rightarrow \psi'(\mathbf{r}, t) = i\gamma^1\gamma^3\psi^*(\mathbf{r}, -t) \quad A_\mu(\mathbf{r}, t) \rightarrow A'_\mu(\mathbf{r}, t) = \tilde{A}_\mu(\mathbf{r}, -t) \quad (5.40)$$

La transformation de A_μ consiste à inverser le signe des composantes spatiales, soit le signe du potentiel vecteur, ce qui revient à changer le signe du champ magnétique sans changer celui du champ électrique. Supposons que les champs transformés respectent encore l'équation de Dirac et vérifions que l'équation originale est récupérée à la fin :

$$\begin{aligned} i\gamma^\mu\partial_\mu\psi'(\mathbf{r}, t) - e\gamma^\mu A'_\mu(\mathbf{r}, t)\psi'(\mathbf{r}, t) - m\psi'(\mathbf{r}, t) &= 0 \\ &= i\gamma^\mu\gamma^1\gamma^3\partial_\mu\psi^*(\mathbf{r}, -t) - e\gamma^\mu\tilde{A}_\mu(\mathbf{r}, -t)\gamma^1\gamma^3\psi^*(\mathbf{r}, -t) - m\gamma^1\gamma^3\psi^*(\mathbf{r}, -t) \\ &= i\gamma^3\gamma^1\gamma^\mu\gamma^1\gamma^3\partial_\mu\psi^*(\mathbf{r}, -t) - e\gamma^3\gamma^1\gamma^\mu\tilde{A}_\mu(\mathbf{r}, -t)\gamma^1\gamma^3\psi^*(\mathbf{r}, -t) - m\psi^*(\mathbf{r}, -t) \\ &= -i\gamma^3\gamma^1\gamma^\mu\gamma^1\gamma^3\tilde{\partial}_\mu\psi^*(\mathbf{r}, t) - e\gamma^3\gamma^1\gamma^\mu\tilde{A}_\mu(\mathbf{r}, t)\gamma^1\gamma^3\psi^*(\mathbf{r}, t) - m\psi^*(\mathbf{r}, t) \end{aligned} \quad (5.41)$$

À la deuxième ligne nous avons omis le facteur i dans ψ' puisqu'il apparaît partout. À la troisième ligne nous avons multiplié par $\gamma^3\gamma^1$ et à la ligne suivante nous avons changé t en $-t$, ce qui change le signe des dérivées temporelles, donc $\partial_\mu \rightarrow -\tilde{\partial}_\mu$. Appliquons ensuite une conjugaison complexe, sachant que seul γ^2 est imaginaire, les autres matrices étant réelles :

$$i\gamma^3\gamma^1\gamma^{\mu*}\gamma^1\gamma^3\tilde{\partial}_\mu\psi(\mathbf{r}, t) - e\gamma^3\gamma^1\gamma^{\mu*}\gamma^1\gamma^3\tilde{A}_\mu(\mathbf{r}, t)\psi - m\psi = 0 \quad (5.42)$$

On vérifie sans peine que

$$\gamma^3\gamma^1\gamma^{0*}\gamma^1\gamma^3 = \gamma^0 \quad \gamma^3\gamma^1\gamma^{1,2,3*}\gamma^1\gamma^3 = -\gamma^{1,2,3} \quad (5.43)$$

ou encore $\gamma^3 \gamma^1 \gamma^{\mu*} \gamma^1 \gamma^3 = \tilde{\gamma}^\mu$. Au total, cela nous permet d'écrire

$$i \tilde{\gamma}^\mu \tilde{\partial}_\mu \psi(\mathbf{r}, t) - e \tilde{\gamma}^\mu \tilde{A}_\mu(\mathbf{r}, t) \psi - m \psi = 0 \quad (5.44)$$

ce qui revient à

$$i \gamma^\mu \partial_\mu \psi(\mathbf{r}, t) - e \gamma^\mu A_\mu(\mathbf{r}, t) \psi - m \psi = 0 \quad (5.45)$$

soit l'équation de Dirac originale.

Inversion du temps en mécanique quantique Considérons l'équation de Schrödinger générale :

$$i \frac{\partial}{\partial t} |\psi\rangle = H |\psi\rangle \quad (5.46)$$

En appliquant une conjugaison complexe en même temps que la transformation $t \mapsto -t$, on obtient

$$i \frac{\partial}{\partial t} |\psi\rangle^* = H^* |\psi\rangle^* \quad (5.47)$$

L'équation de Schrödinger est encore valable pour le nouvel état $|\psi\rangle^*$, pourvu que $H = H^*$. L'invariance par inversion du temps de l'équation de Schrödinger impose donc une condition au hamiltonien, qui n'est pas toujours respectée pour un hamiltonien non fondamental. Par exemple, l'imposition du champ magnétique statique brise l'invariance par inversion du temps.

T est antilinéaire L'opération d'inversion du temps n'est pas représentée par un opérateur, comme Π ou C , avec des valeurs propres bien définies, même si le carré de cette opération est manifestement l'unité. La raison en est que, d'après Wigner, l'opération d'inversion du temps est représentée par un opérateur *anti-linéaire*, c'est-à-dire un opérateur T tel que

$$T \lambda |\psi\rangle = \lambda^* T |\psi\rangle \quad (5.48)$$

Un opérateur anti-linéaire qui préserve la norme des états est qualifié d'*antiunitaire*. Un opérateur anti-linéaire ne peut pas vraiment être représenté par une matrice, et la notion de valeur propre n'a pas de sens dans ce cas. Par contre, un opérateur A devient $T A T^{-1}$ par inversion du temps.

Comment voir que l'opérateur T doit être antilinéaire? Considérons l'opérateur d'évolution e^{-iHt} , dont l'action sur un état est $e^{-iHt} |\Psi(0)\rangle = |\Psi(t)\rangle$. Appliquer l'opérateur d'évolution e^{iHt} dans le sens contraire du temps sur l'état $T |\Psi(0)\rangle$, et ensuite appliquer de nouveau l'inversion du temps T , devrait donner la même chose que d'évoluer l'état original dans le sens direct du temps, en supposant que le hamiltonien soit lui-même invariant par inversion du temps :

$$T e^{iHt} T |\Psi(0)\rangle = e^{-iHt} |\Psi(0)\rangle \quad (5.49)$$

En développant cette expression pour des temps petits, on trouve

$$T i H T = -i H \quad (5.50)$$

Si T était un opérateur unitaire, on pourrait diviser par i et conclure que $T H T = -H$ et donc que $T H = -H T$ (car alors $T^2 = 1$). Mais dans ce cas, on pourrait conclure, à tort, qu'à tout état propre

$|\Psi\rangle$ du hamiltonien, d'énergie E , correspond un autre état, $T|\Psi\rangle$, d'énergie $-E$. En effet, $TH|\Psi\rangle = TE|\Psi\rangle = -HT|\Psi\rangle$ donc $HT|\Psi\rangle = -ET|\Psi\rangle$. Si, au contraire, l'opérateur est antiunitaire, on a plutôt

$$TiHT = -iTHT = -iH \implies [T, H] = 0 \quad (5.51)$$

Cependant, un opérateur antilinéaire ne possède pas les propriétés habituelles d'un opérateur linéaire, en particulier l'existence de vecteurs et valeurs propres. En effet, si on pose $T|\Psi\rangle = \lambda|\Psi\rangle$, alors $Ti|\Psi\rangle = -i\lambda|\Psi\rangle$, de sorte que la notion de sous-espace propre et de valeur propre n'a pas de sens pour un tel opérateur.

Carré de l'opérateur T Le carré de l'opérateur T n'est pas nécessairement l'unité. En particulier, pour une particule de spin $\frac{1}{2}$, on montre que $T^2 = -1$. Expliquons; le fait que l'opérateur du spin change de signe lors de l'inversion du temps peut s'écrire $T\mathbf{S} = -\mathbf{S}T$. Autrement dit, mesurer le moment cinétique sur un état inversé dans le temps devrait produire l'opposé de la valeur obtenue sur l'état original. Nous allons vérifier que cette propriété est vérifiée si l'action de T sur les états propres du spin $|\uparrow\rangle$ et $|\downarrow\rangle$ est

$$T|\uparrow\rangle = i|\downarrow\rangle \quad \text{et} \quad T|\downarrow\rangle = -i|\uparrow\rangle \quad (5.52)$$

Calculons maintenant l'effet de T sur $S_z|\uparrow\rangle$ et $S_z|\downarrow\rangle$:

$$TS_z|\uparrow\rangle = T\frac{1}{2}|\uparrow\rangle = \frac{i}{2}|\downarrow\rangle \quad \text{et} \quad TS_z|\downarrow\rangle = -T\frac{1}{2}|\downarrow\rangle = \frac{i}{2}|\uparrow\rangle \quad (5.53)$$

alors que, dans l'ordre inverse,

$$S_zT|\uparrow\rangle = S_z i|\downarrow\rangle = -\frac{i}{2}|\downarrow\rangle \quad \text{et} \quad S_zT|\downarrow\rangle = -S_z i|\uparrow\rangle = -\frac{i}{2}|\uparrow\rangle \quad (5.54)$$

On constate que le résultat de l'opération est opposé, comme attendu. Pour S_x , on trouve

$$TS_x|\uparrow\rangle = T\frac{1}{2}|\downarrow\rangle = -\frac{i}{2}|\uparrow\rangle \quad \text{et} \quad TS_x|\downarrow\rangle = T\frac{1}{2}|\uparrow\rangle = \frac{i}{2}|\downarrow\rangle \quad (5.55)$$

et, dans l'ordre inverse,

$$S_xT|\uparrow\rangle = S_x i|\downarrow\rangle = \frac{i}{2}|\uparrow\rangle \quad \text{et} \quad S_xT|\downarrow\rangle = -S_x i|\uparrow\rangle = -\frac{i}{2}|\downarrow\rangle \quad (5.56)$$

Encore une fois, les résultats sont opposés. Finissons par S_y :

$$TS_y|\uparrow\rangle = T(-\frac{i}{2}|\downarrow\rangle) = \frac{i}{2}T|\downarrow\rangle = -\frac{1}{2}|\uparrow\rangle \quad \text{et} \quad TS_y|\downarrow\rangle = T(\frac{i}{2}|\uparrow\rangle) = -\frac{i}{2}T|\uparrow\rangle = \frac{1}{2}|\downarrow\rangle \quad (5.57)$$

et, dans l'ordre inverse,

$$S_yT|\uparrow\rangle = -iS_y|\downarrow\rangle = \frac{1}{2}|\uparrow\rangle \quad \text{et} \quad S_yT|\downarrow\rangle = iS_y|\uparrow\rangle = -\frac{1}{2}|\downarrow\rangle \quad (5.58)$$

ce qui est encore une fois le résultat attendu. Donc notre hypothèse sur l'action de T dans cette base est vérifiée. Par contre, il est clair que $T^2 = -1$. Cette conclusion est générale, pour tous les systèmes comportant des spin demi-entiers.

Dégénérescence de Kramers Une des conséquences de l'invariance par inversion du temps est la dégénérescence de Kramers : dans tout système comportant un moment cinétique demi-entier, les états propres du hamiltonien sont doublement dégénérés. En effet, soit $|\Psi\rangle$ un état propre de H avec valeur propre E . L'état $T|\Psi\rangle$ est aussi un état propre de H avec la même valeur propre, car $[H, T] = 0$. Par contre, si cet état était non dégénéré, il faudrait que $T|\Psi\rangle = \alpha|\Psi\rangle$ et alors $T^2|\Psi\rangle = |\alpha|^2|\Psi\rangle$, ce qui est impossible si $T^2 = -1$. Donc $T|\Psi\rangle$ n'est pas proportionnel à $|\Psi\rangle$ et le niveau E est au moins doublement dégénéré.

Bilan détaillé Une conséquence immédiate de la symétrie d'inversion du temps est le principe du *bilan détaillé* ou encore de *micro réversibilité*, qui stipule qu'une réaction qui se produit dans un sens ($A+B \rightarrow C+D$) doit aussi se produire dans l'autre sens ($C+D \rightarrow A+B$) pour des conditions identiques d'énergie, si on échange les états initiaux et finaux. C'est ce qui est observé dans le cadre des interactions électromagnétiques et fortes; seules les interactions faibles pourraient en principe briser (très faiblement) ce principe, mais leur faiblesse empêche toute confirmation expérimentale de cette violation de la micro réversibilité.

Absence de dipôle électrique permanent Une autre conséquence de la symétrie T est qu'une particule subatomique ne doit pas posséder de moment dipolaire électrique permanent si elle possède un spin. L'argument est le suivant : l'opérateur \mathbf{d} du dipôle électrique est un vecteur polaire qui ne change pas de signe lors de l'inversion du temps :

$$\Pi \mathbf{d} \Pi = -\mathbf{d} \quad T \mathbf{d} T^{-1} = \mathbf{d} , \quad (5.59)$$

ce qui est exactement le contraire de l'opérateur du spin :

$$\Pi \mathbf{S} \Pi = \mathbf{S} \quad T \mathbf{S} T^{-1} = -\mathbf{S} , \quad (5.60)$$

(notons que $\Pi = \Pi^{-1}$ alors que $T^{-1} = \pm T$, selon le système étudié). Si $|\Psi\rangle$ est l'état fondamental de la particule, alors le moment dipolaire électrique de cette particule est $\langle \mathbf{d} \rangle = \langle \Psi | \mathbf{d} | \Psi \rangle$. Appliquons la relation ci-haut pour la parité :

$$\langle \mathbf{d} \rangle = \langle \Psi | \mathbf{d} | \Psi \rangle = -\langle \Psi | \Pi \mathbf{d} \Pi | \Psi \rangle \quad (5.61)$$

Si la parité est conservée, alors $\Pi|\Psi\rangle = \pm|\Psi\rangle$ et cette dernière expression devient $-\langle \Psi | \mathbf{d} | \Psi \rangle$, d'où on conclut que $\langle \mathbf{d} \rangle = 0$: le dipôle électrique permanent est nul, conséquence de l'invariance par inversion. Cette conclusion ne s'applique pas au spin $\langle \mathbf{S} \rangle$, qui est un vecteur axial. Comme la symétrie P est violée dans les interactions faibles de manière importante, cet argument ne suffit pas à interdire la présence d'un moment dipolaire électrique permanent. Considérons alors le même argument, cette fois en utilisant la symétrie T :

$$\langle \mathbf{d} \rangle = \langle \Psi | \mathbf{d} | \Psi \rangle = \langle \Psi | T \mathbf{d} T^{-1} | \Psi \rangle \quad (5.62)$$

La relation correspondante pour le spin est

$$\langle \mathbf{S} \rangle = \langle \Psi | \mathbf{S} | \Psi \rangle = -\langle \Psi | T \mathbf{S} T^{-1} | \Psi \rangle \quad (5.63)$$

Or, si le spin est non nul et que T est conservé, la relation ci-dessus entraîne que $T^{-1}|\Psi\rangle = \alpha|\Psi\rangle$ et $\langle \Psi | T = \beta \langle \Psi |$ avec des constantes α et β telles que $\alpha\beta = -1$. Les mêmes constantes dans l'équation précédente mènent à la conclusion que $\langle \mathbf{d} \rangle = -\langle \mathbf{d} \rangle$, ce qui conclut la preuve que $\langle \mathbf{d} \rangle = 0$. L'expérience la plus précise à ce sujet donne la limite supérieure suivante au moment dipolaire électrique du neutron : $|\mathbf{d}|/e < 6 \times 10^{-27} \text{ m}$.

Théorème CPT Un théorème célèbre, attribué à SCHWINGER, LÜDERS et PAULI, stipule que pour toute théorie des champs invariante de Lorentz, l'action consécutive de C , P et T doit laisser la théorie invariante, même si l'une des trois opérations n'est pas une symétrie de la théorie. Autrement dit, si on échange particules et antiparticules, qu'on reflète l'Univers dans un miroir et qu'on inverse le sens du temps, on devrait observer les mêmes processus avec les mêmes probabilités.

Une conséquence du théorème CPT est que la masse d'une particule doit être strictement égale à celle de son antiparticule. Ceci est vérifié de manière très précise dans les systèmes de kaons neutres, où la différence relative de masse entre le K^0 et le \bar{K}^0 est inférieure à 6×10^{-19} .

Nous verrons que le modèle standard comporte une violation faible de la combinaison CP , ce qui revient alors à une violation de la symétrie T . Cette violation est importante pour expliquer l'asymétrie matière-antimatière dans l'Univers. Elle peut aussi permettre l'existence d'un moment dipolaire électrique pour le neutron, mais à la hauteur de $|\mathbf{d}/e| \sim 10^{-31} \text{m}$ seulement, ce qui est beaucoup plus petit que la limite observée actuellement.

B Isospin

5.B.1 Le nucléon

En 1932, tout juste après la découverte du neutron, HEISENBERG introduisit une idée très fructueuse : ayant remarqué que le proton et le neutron ont presque la même masse et qu'ils semblent se comporter de manière identique face aux interactions fortes, il émit l'hypothèse que ces deux particules ne sont que deux états différents d'une même particule qu'il appela *nucléon*. Le nucléon serait doté d'une propriété qu'il appela *isospin*, semblable au spin ordinaire, pouvant prendre deux valeurs : $\frac{1}{2}$ et $-\frac{1}{2}$. Le proton serait l'état d'isospin $+\frac{1}{2}$ et le neutron l'état d'isospin $-\frac{1}{2}$. On pourrait schématiquement représenter le nucléon comme un doublet comportant deux composantes :

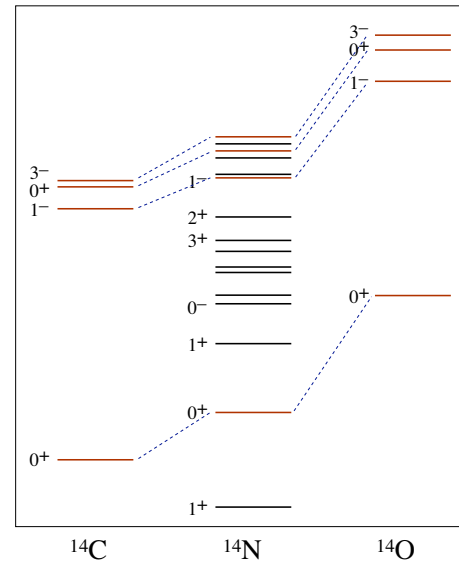
$$N = \begin{pmatrix} p \\ n \end{pmatrix} \quad (5.64)$$

Par analogie avec le spin ordinaire, l'isospin forme un ensemble de trois opérateurs (I_1, I_2, I_3) qui ont les mêmes relations de commutation que les composantes du moment cinétique, sauf qu'ils forment un vecteur dans un espace abstrait, dit « espace d'isospin », et non dans l'espace physique à trois dimensions. Le fait que le proton et le neutron ne soient pas entièrement équivalents ne serait dû qu'à la différence de charge électrique entre les deux, de la même manière que les deux orientations du spin d'un électron n'ont pas la même énergie en présence d'un champ magnétique. Autrement dit, l'interaction électromagnétique mène en quelque sorte à un « effet Zeeman » de l'isospin. Mais si on fait abstraction de cet effet, les phénomènes n'impliquant que l'interaction forte seraient invariants lorsqu'on procède à une rotation dans l'espace d'isospin.

Ceci est un point important : de même qu'une rotation dans l'espace ordinaire doit être accompagnée d'une combinaison des deux projections de spin (voir la section 3.A), une rotation dans l'espace de l'isospin induit une superposition du proton et du neutron. Un état général du nucléon est donc une combinaison de neutron et de proton; seule une mesure de la propriété I_3 nous permet de savoir si le nucléon est un neutron ou un proton. De plus, les isospins de plusieurs nucléons s'additionnent comme le moment cinétique : deux nucléons forment ainsi un système qui possède un triplet d'isospin $I = 1$ et un singulet d'isospin $I = 0$. L'état triplet est symétrique lorsque les deux nucléons sont échangés, alors que l'état singulet est antisymétrique. Les différentes projections I_3 du triplet auront en fait des énergies différentes en raison de l'interaction électromagnétique et de la différence de masse entre le neutron et le proton. Voir la figure 5.1 pour un exemple tiré de la physique nucléaire.

FIGURE 5.1

États de basse énergie de trois isobares. Les états analogues forment un triplet sont indiqués en rouge et sont reliés par des lignes pointillées. Le noyau $^{14}_7\text{N}$ comporte aussi plusieurs états qui sont des singulets d'isospin. Les membres d'un même multiplet ont les mêmes valeurs du spin J et de la parité.



Le nucléon n'est pas la seule particule à posséder un isospin. Pratiquement tous les hadrons se sont fait assigner des valeurs d'isospin, au fur et à mesure qu'ils ont été découverts. Les hadrons peuvent donc être groupés en multiplets dont les membres ont des charges différentes, des masses semblables, mais dont les autres propriétés (spin, parité) sont identiques. Par exemple, les trois pions (π^+ , π^0 et π^-) forment un triplet d'isospin $I = 1$ et de spin-parité 0^- . Les quatre particules Delta (Δ^{++} , Δ^+ , Δ^0 et Δ^-) forment un quadruplet d'isospin $I = \frac{3}{2}$, de spin-parité $\frac{3}{2}^+$ et de masse $m_\Delta \sim 1232 \text{ MeV}$.

5.B.2 Représentations de l'isospin

La théorie de l'isospin est formellement identique à celle du moment cinétique. Les trois opérateurs de l'isospin $I_{1,2,3}$ ont exactement les mêmes relations de commutation que les composantes du moment cinétique. On définit également le carré de l'isospin $\mathbf{I}^2 = I_1^2 + I_2^2 + I_3^2$ qui joue le même rôle que le carré du moment cinétique \mathbf{J}^2 , c'est-à-dire qu'il commute avec les trois opérateurs I_a . Rappelons le résultat central de la théorie du moment cinétique dans ce contexte.

Considérons un sous-espace de Hilbert V de dimension d dans lequel l'action des opérateurs I_a est

bien définie. Cela signifie que, dans V , les opérateurs I_a ont une forme matricielle $d \times d$ et que ces matrices obéissent aux relations de commutation $[I_a, I_b] = 2i\epsilon_{abc}I_c$. La dimension de V est alors $d = 2I + 1$, où I peut prendre des valeurs positives ou nulles espacées de $\frac{1}{2}$: $I = 0, \frac{1}{2}, 1, \frac{3}{2}, \dots$. Pour une valeur donnée de I , l'opérateur I^2 est $I(I + 1)$ fois la matrice identité. Les $2I + 1$ états de base de V peuvent être choisis comme les états propres de I_3 , dont les valeurs propres s'échelonnent de $-I$ à $+I$, espacées de 1. Les deux autres opérateurs (I_1 et I_2) ne sont pas diagonaux dans cette base, car ils ne commutent pas avec I_3 , malgré qu'ils aient exactement le même spectre de valeurs propres. Notons $|m\rangle$ un état propre de I_3 avec valeur propre m . On peut définir les opérateurs d'échelle

$$I_+ = I_1 + iI_2 \quad \text{et} \quad I_- = I_1^\dagger = I_1 - iI_2 \quad (5.65)$$

tels que

$$I_+|m\rangle = \sqrt{I(I+1) - m(m+1)}|m+1\rangle \quad \text{et} \quad I_-|m\rangle = \sqrt{I(I+1) - m(m-1)}|m-1\rangle \quad (5.66)$$

L'action des opérateurs I_\pm s'arrête aux extrémités du spectre des valeurs propres :

$$I_+|m=I\rangle = 0 \quad \text{et} \quad I_-|m=-I\rangle = 0 \quad (5.67)$$

L'action de I_\pm nous permet de retrouver une représentation explicite des matrices I_a dans V :

$$\begin{aligned} \langle m'|I_3|m\rangle &= m\delta_{m',m} \\ \langle m'|I_+|m\rangle &= \delta_{m',m+1}\sqrt{j(j+1) - m(m+1)} \\ \langle m'|I_-|m\rangle &= \delta_{m',m-1}\sqrt{j(j+1) - m(m-1)} \end{aligned} \quad (5.68)$$

Par exemple, dans le cas $I = 1$, on trouve

$$I_1 = \frac{1}{\sqrt{2}} \begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 1 \\ 0 & 1 & 0 \end{pmatrix} \quad I_2 = \frac{1}{\sqrt{2}} \begin{pmatrix} 0 & -i & 0 \\ i & 0 & -i \\ 0 & i & 0 \end{pmatrix} \quad I_3 = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & -1 \end{pmatrix} \quad (5.69)$$

alors que dans le cas $I = \frac{1}{2}$, on trouve simplement les matrices de Pauli (divisées par deux) :

$$I_1 = \frac{1}{2} \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \quad I_2 = \frac{1}{2} \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix} \quad I_3 = \frac{1}{2} \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \quad (5.70)$$

L'espace V de dimension $2I + 1$ généré par les états propres de I_3 constitue ce qu'on appelle communément une *représentation* de l'isospin (ou du moment cinétique, dans un autre contexte). Plus précisément, le sous-espace est appelé le *module* de la représentation, et l'expression explicite des matrices I_a constitue ce qu'on appelle la *représentation* proprement dite. Mais sur ce point la terminologie est un peu lâche et le mot « représentation » est appliqué indifféremment à la forme des matrices ou au sous-espace lui-même. On peut affirmer que le but de la théorie du moment cinétique est de trouver la forme de toutes les représentations possibles du moment cinétique, ce que nous avons résumé ici dans le contexte de l'isospin.

5.B.3 Théorème de Noether

Voyons comment la symétrie d'isospin se manifeste dans le contexte d'une théorie définie par un lagrangien. Considérons deux champs de Dirac, ψ_1 et ψ_2 , représentant des espèces de particules différentes, par exemple le proton et le neutron. On suppose cependant que ces deux particules ont exactement la même masse. Le lagrangien de cette théorie est

$$\mathcal{L} = \sum_{j=1,2} \{ i \bar{\psi}_j \not{\partial} \psi_j - m \bar{\psi}_j \psi_j \} \quad (5.71)$$

Ce lagrangien est invariant lorsqu'on procède à une rotation unitaire des deux champs : $\psi_j \rightarrow \psi'_j = U_{jk} \psi_k$ (sommé sur k) où U est une matrice unitaire d'ordre 2. En effet, le spineur conjugué se transforme alors comme $\bar{\psi}'_j = U_{jk}^* \bar{\psi}_k$ et donc, par exemple, l'expression $\bar{\psi}_j \psi_j$ (sommée sur j) est invariante :

$$\bar{\psi}_j \psi_j \rightarrow U_{jk} U_{lj}^* \bar{\psi}_l \psi_k \quad \text{mais} \quad U_{kj} U_{lj}^* = (U U^\dagger)_{kl} = \delta_{kl} \quad (5.72)$$

en vertu de l'unitarité de la matrice U , ce qui nous ramène à l'expression d'origine. Donc le lagrangien jouit d'une symétrie par toute transformation unitaire appartenant au groupe $U(2)$.

L'existence d'une symétrie continue se traduit par une quantité conservée. Ceci est une conséquence du *théorème de Noether*, que nous allons maintenant énoncer, puis démontrer pour des symétries particulièrement simples comme la transformation unitaire définie ci-haut. Considérons une théorie exprimée en fonction d'un champ ou d'une collection de champs qu'on désigne par ψ . Par exemple, ψ pourrait être un doublet de champs de Dirac (ψ_1, ψ_2) , comme introduit dans le contexte de l'isospin; dans ce cas, ψ possède deux indices $(\psi_{\alpha i})$: un indice de spineur α prenant quatre valeurs et un indice d'isospin i prenant deux valeurs. Nous allons supposer que la densité lagrangienne \mathcal{L} , qui est une fonction de ψ et de ses dérivées $\partial_\mu \psi$, est invariante lors d'une certaine transformation continue

$$\psi(x) \rightarrow \psi'(x) = \mathcal{F}(\psi)(x) \quad (5.73)$$

et que cette transformation n'affecte pas les coordonnées d'espace-temps x (le théorème de Noether s'applique aussi au cas de transformations qui affectent les coordonnées, mais nous n'aborderons pas cet aspect). Par exemple, la transformation unitaire introduite plus haut correspond à

$$\mathcal{F}(\psi) = U \psi \quad (5.74)$$

où U est une matrice unitaire d'ordre 2 agissant sur le doublet $\psi = (\psi_1, \psi_2)$. Comme la transformation \mathcal{F} est continue, elle existe en version infinitésimale, c'est-à-dire qu'on peut considérer une transformation arbitrairement proche de l'identité et la noter comme suit :

$$\psi \rightarrow \psi' = \psi + i \delta \omega_a T_a \psi \quad (5.75)$$

où $\delta \omega_a$ est un paramètre infinitésimal, T_a représente une certaine opération linéaire sur ψ et a est un indice représentant différents degrés de liberté dans la transformation. Considérons par exemple une transformation définie par une matrice unitaire d'ordre n , donc appartenant au groupe $U(n)$. Toute matrice unitaire peut être représentée par l'exponentielle complexe d'une matrice hermitienne :

$$U = e^{iH} \quad \text{où} \quad H^\dagger = H \quad \text{car} \quad U^\dagger = e^{-iH} \quad \text{et donc} \quad U^\dagger U = 1 \quad (5.76)$$

Une matrice unitaire arbitrairement proche de la matrice identité peut donc s'écrire

$$U = 1 + i\delta\omega_a T_a \quad (5.77)$$

où les $\delta\omega_a$ sont des paramètres infinitésimaux, et les matrices T_a sont appelées les *générateurs* du groupe $U(n)$ (voir l'annexe 7.D pour plus de détails). Dans le cas de $U(2)$, cette expression devient

$$U = 1 + i\delta\omega_0 \mathbb{I} + i\frac{1}{2} \sum_{a=1}^3 \delta\omega_a \sigma_a \quad \sigma_a : \text{matrices de Pauli} \quad (5.78)$$

Notons à cet effet que toute matrice hermitienne d'ordre deux est une combinaison réelle des trois matrices de Pauli et de la matrice identité. En général, le nombre de matrices hermitiennes d'ordre n et linéairement indépendantes est n^2 , soit exactement le même nombre de paramètres nécessaire pour spécifier une matrice unitaire d'ordre n .

Le théorème de Noether stipule que si l'action est invariante par une transformation (5.74), alors il existe des quantités conservées, décrites par des quadricourants définis par

$$J_a^\mu = -i \frac{\partial \mathcal{L}}{\partial \partial_\mu \psi} T_a \psi \quad (5.79)$$

Ces courants sont conservés, c'est-à-dire qu'ils obéissent à l'équation de continuité $\partial_\mu J_a^\mu = 0$. On peut donc leur associer une charge totale, qui est l'intégrale de la densité :

$$Q_a = \int d^3x J_a^0 \quad (5.80)$$

Dans le cas de la symétrie $U(2)$ décrite ci-dessus, il existe quatre courants conservés :

$$J_0^\mu = \bar{\psi} \gamma^\mu \psi \quad \text{et} \quad J_a^\mu = \bar{\psi} \gamma^\mu \sigma_a \psi \quad (5.81)$$

Le premier (J_0) décrit simplement le courant et la charge électrique (ou simplement le nombre total de particules, les antiparticules comptant négativement), alors que les trois autres correspondent à l'isospin. Si N_1 et N_2 désignent le nombre de particules de type 1 et 2, le nombre total de particules est $N = N_1 + N_2$ et la troisième composante de l'isospin est $I_3 = (N_1 - N_2)/2$. Les deux autres composantes de l'isospin ne s'expriment pas en fonction de N_1 et N_2 directement. En fait, on montre sans difficulté qu'en fonction des opérateurs de création et d'annihilation des fermions, l'opérateur d'isospin a l'expression suivante (on pose $\sigma_0 = \mathbb{I}$) :

$$I_a = \frac{1}{2} \sum_{\mathbf{k}} \sum_{s=1,2} \sum_{j,k} \left\{ c_{j,s,\mathbf{k}}^\dagger (\sigma_a)_{jk} c_{k,s,\mathbf{k}} - d_{j,s,\mathbf{k}}^\dagger (\sigma_a)_{jk} d_{k,s,\mathbf{k}} \right\} \quad (5.82)$$

Preuve du théorème de Noether

Appliquons une transformation infinitésimale (5.75) à l'action. Supposons en outre que les paramètres $\delta\omega_a$ ne sont pas constants, mais dépendent de la position et du temps d'une manière quelconque. La variation correspondante de l'action est :

$$\delta S = \int d^4x \left\{ \frac{\partial \mathcal{L}}{\partial \psi} \delta \psi + \frac{\partial \mathcal{L}}{\partial \partial_\mu \psi} \delta(\partial_\mu \psi) \right\} \quad (5.83)$$

où

$$\delta\psi = i\delta\omega_a T_a \psi \quad \text{et} \quad \partial_\mu \delta\psi = i\partial_\mu \delta\omega_a T_a \psi + i\delta\omega_a T_a \partial_\mu \psi \quad (5.84)$$

Mais l'action est invariante par rapport à une transformation de type (5.75), pourvu que les paramètres $\delta\omega_a$ soient constants. Autrement dit, la symétrie en question est *globale* (la même transformation partout), et non *locale* (indépendante d'une position à l'autre). Donc tous les termes qui n'impliquent pas $\partial_\mu \delta\omega_a$ dans δS doivent s'annuler pour cette raison. Il reste donc

$$\delta S = i \int d^4x \frac{\partial \mathcal{L}}{\partial \partial_\mu \psi} T_a \psi \partial_\mu \delta\omega_a(x) \quad (5.85)$$

En intégrant cette variation par parties, on obtient

$$\delta S = -i \int d^4x \partial_\mu \left\{ \frac{\partial \mathcal{L}}{\partial \partial_\mu \psi} T_a \psi \right\} \delta\omega_a(x) \quad (5.86)$$

On utilise ensuite le principe de moindre action qui stipule que la variation δS doit s'annuler pour toute variation du champ, si ce dernier obéit aux équations du mouvement. Ceci entraîne que l'expression qui multiplie $\delta\omega_a(x)$ doit s'annuler identiquement :

$$-i\partial_\mu \left\{ \frac{\partial \mathcal{L}}{\partial \partial_\mu \psi} T_a \psi \right\} = 0 \quad \text{ou} \quad \partial_\mu J_a^\mu = 0 \quad \text{avec} \quad J_a^\mu = -i \frac{\partial \mathcal{L}}{\partial \partial_\mu \psi} T_a \psi \quad (5.87)$$

ce qui démontre le théorème de Noether.

U(2) vs SU(2) La symétrie $U(2)$ mène à l'existence de 4 quantités conservées, qui sont le nombre de particules N et les trois composantes I_a de l'isospin. Il se trouve que la conservation de N est indépendante de celle de l'isospin. Par cela on veut dire que N commute avec les trois opérateurs $I_{1,2,3}$, de même que la matrice unitaire \mathbb{I} commute avec les trois matrices de Pauli. Cela a pour conséquence que les valeurs propres de N ne sont pas affectées par les opérateurs I_a (en particulier les opérateurs d'échelle I_\pm) et qu'à l'intérieur d'un espace V associé à une valeur donnée de \mathbb{I}^2 , la valeur propre de N doit être constante. À la rigueur, on peut avoir plusieurs copies de cet espace avec des valeurs propres différentes de N .

La raison mathématique de cet état de fait est que le groupe $U(2)$ n'est pas *simple*, mais est plutôt le produit de deux groupes simples : $U(2) = SU(2) \times U(1)$. Un groupe G est dit *simple* si ses seuls sous-groupes normaux sont le groupe entier et l'élément identité. Un sous-groupe H est dit *normal* si chacun de ses éléments h est transformé en un autre de ses éléments par tout élément g de G : $ghg^{-1} \in H$. En pratique, cette définition signifie que les générateurs du sous-groupe normal commutent avec ceux du groupe complet.

Étudions la question en fonction d'une représentation exponentielle des éléments du groupe. Nous avons vu que toute matrice unitaire U peut s'exprimer comme

$$U = e^{i\omega_a T_a} \quad (5.88)$$

où les n paramètres ω_a servent à spécifier un élément particulier du groupe, alors que les matrices hermitiennes T_a sont les *générateurs* du groupe (voir annexe 7.D). Les relations de commutations

des matrices T_a déterminent à elles seules la structure du groupe. En général, on peut les écrire comme

$$[T_a, T_b] = i f_{abc} T_c \quad (\text{sommé sur } c) \quad (5.89)$$

où les constantes réelles f_{abc} sont appelées *constantes de structure*. Dans le cas du groupe $U(2)$, on a

$$f_{abc} = \varepsilon_{abc} \quad \text{si } a, b, c \neq 0 \quad \text{et} \quad f_{abc} = 0 \quad \text{si l'un des indices est } 0 \quad (5.90)$$

L'opérateur I_0 commute donc complètement avec les trois autres et son spectre (l'ensemble de ses valeurs propres) est déterminé de manière indépendante de celui de $I_{1,2,3}$.

L'opérateur I_0 est en fait associé aux transformations de phase pure $\psi' = e^{i\theta} \psi$. Ces transformations commutent avec les transformations unitaires $\psi' = U\psi$ définies par des matrices de déterminant unité (et donc appartenant à $SU(2)$). La théorie définie par le lagrangien (5.71) serait invariante par une transformation de phase pure même si les masses des deux espèces de fermions ψ_1 et ψ_2 étaient différentes et n'a donc rien à voir avec la symétrie particulière qui nous intéresse ici, qui est liée à l'égalité des masses.

C SU(3) et organisation des hadrons

5.C.1 Étrangeté

À partir de 1947, on découvrit des particules d'un type nouveau, qu'on qualifia de « particules étranges ». Ce qualificatif est dû à leur temps de vie relativement long, qui contraste avec une section efficace de production relativement grande. Autrement dit, ces particules sont produites en grand nombre lorsque des pions ou des protons frappent une cible – ce qui semble indiquer que l'interaction forte est responsable de la production – alors que leur long temps de vie semble indiquer que leur désintégration est causée par l'interaction faible seulement. La première particule étrange découverte fut plus tard appelée *kaon*, et existe en quatre variétés : K^+ , K^0 , \bar{K}^0 et K^- , avec environ trois fois la masse du pion. Il y eut aussi le « lambda » (Λ), particule neutre de spin $\frac{1}{2}$ et de masse 1 116 MeV, manifestement un baryon comme le nucléon.

A. Pais suggéra le premier d'expliquer la stabilité de ces particules en supposant qu'elles sont produites par l'interaction forte, mais par paires et avec des propriétés opposées ; elles se désintègrent ensuite par interaction faible, donc beaucoup plus lentement. M. GELL-MANN et K. NISHIJIMA supposèrent l'existence d'un nombre quantique appelé *étrangeté* et noté S , qui prendrait des valeurs entières. De manière arbitraire, on assigna une étrangeté $S = 1$ au K^+ et au K^0 , ce qui impose la valeur $S = -1$ au K^- et au \bar{K}^0 , car les particules d'étrangetés opposées sont produites ensemble.

Les hadrons connus furent organisés avec succès par Gell-Mann en 1961 (*the eightfold way*). Cette classification le mena ensuite à postuler que les hadrons sont constitués d'unités primitives qu'il appela *quarks* (d'après un passage du roman *Finnegan's wake* de James Joyce). La production associée est simple à comprendre à l'aide du modèle des quarks : les particules étranges contiennent un

quark s ou un antiquark \bar{s} . Lorsqu'un faisceau de pions ou de protons frappe une cible, des paires $s\bar{s}$ sont créées copieusement, car il s'agit d'un processus gouverné par l'interaction forte. Cependant, le s et le \bar{s} sont séparés et se retrouvent dans des hadrons différents, ceux contenant le quark s ayant une étrangeté $S = -1$ et ceux héritant du \bar{s} ayant une étrangeté $S = +1$. Ce phénomène porte le nom de *production associée*

5.C.2 $SU(3)$

Voyons comment cette classification découle de l'existence des quarks. Supposons pour commencer l'existence de trois espèces de particules élémentaires, décrites par l'équation de Dirac, qu'on désignera par les symboles u , d et s . On pourrait également noter les champs correspondants $\psi_{1,2,3}$, comme lors de notre discussion de l'isospin, mais cette fois avec trois champs au lieu de deux. Le lagrangien décrivant cette théorie est

$$\mathcal{L} = \sum_{j=1}^3 \{ i\bar{\psi}_j \not{\partial} \psi_j - m\bar{\psi}_j \psi_j \} \quad (5.91)$$

et jouit d'une symétrie $U(3)$, car l'action est invariante par toute rotation unitaire des trois champs $\psi_{1,2,3}$. Ce groupe est cependant le produit de deux groupes simples : $U(3) = SU(3) \times U(1)$. Comme dans le cas de l'isospin, la symétrie pertinente ici correspond au groupe $SU(3)$ seulement. Comme ce groupe est de dimension $3^2 - 1 = 8$, il compte huit générateurs T_a (voir annexe 7.D).

Se pose maintenant le même problème que dans la théorie de l'isospin (ou du moment cinétique), à savoir de déterminer quelles sont toutes les représentations possibles de $SU(3)$. Autrement dit, quels sont les modules V dans lesquels les huit opérateurs T_a ont une forme matricielle bien définie respectant les relations de commutation ci-haut. Comme pour l'isospin, il s'agit d'un problème fondamental de la théorie des groupes de Lie. Nous ne pourrions dans le contexte de ce cours démontrer la solution générale à ce problème ; nous nous contenterons d'en donner la solution générale sans démonstration et de l'illustrer avec des exemples simples.

Remarquons d'abord que deux des générateurs de $SU(3)$ commutent entre eux : T_3 et T_8 . On peut donc désigner les états propres d'une représentation par les valeurs propres respectives de ces deux opérateurs, notées m et y :

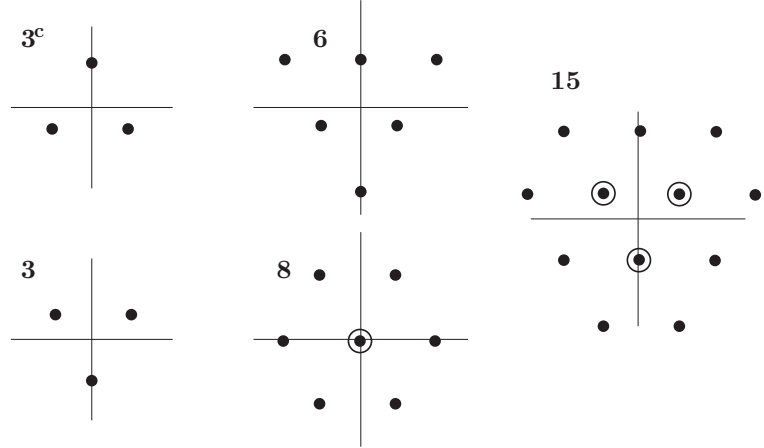
$$T_3|m, y\rangle = m|m, y\rangle \quad T_8|m, y\rangle = y|m, y\rangle \quad (5.92)$$

Notons que le générateur T_3 de $SU(3)$ est souvent noté I_3 , et on introduit aussi l'opérateur d'hypercharge $Y = (2/\sqrt{3})T_8$.

Représentation triviale La représentation la plus simple de $SU(3)$ est la représentation triviale, ou le *singulet*, ne comportant qu'un seul état $|0\rangle$, qui n'est nullement affecté par aucun des opérateurs : $T_a|0\rangle = |0\rangle$.

FIGURE 5.2

Quelques représentations de $SU(3)$, représentées sur le plan (m, y) des valeurs propres de T_3 et T_8 . Un point encerclé désigne deux états dégénérés.



Représentation fondamentale La représentation dite *fondamentale* de $SU(3)$ est de dimension trois et est définie par les matrices de Gell-Mann elles-mêmes (divisées par 2). On peut lire les valeurs propres directement de la forme explicite de λ_3 et λ_8 :

$$(m, y) = \left(\frac{1}{2}, \frac{1}{2\sqrt{3}}\right), \left(-\frac{1}{2}, \frac{1}{2\sqrt{3}}\right), \left(0, -\frac{1}{\sqrt{3}}\right) \quad (5.93)$$

En portant ces valeurs sur un plan cartésien, on forme un triangle équilatéral inversé; voir la figure 5.2. Cette représentation fondamentale est aussi désignée par la lettre **3** (un chiffre en caractères gras indiquant la dimension de la représentation).

Représentation fondamentale conjuguée Supposons que nous ayons une certaine représentation des générateurs T_a , de dimension quelconque, et obéissant donc aux relations de commutation $[T_a, T_b] = if_{abc} T_c$. La représentation conjuguée à celle-ci est définie en appliquant la conjugaison complexe et un changement de signe aux générateurs : $T_a^{(c)} = -T_a^*$. Ces opérateurs conjugués forment aussi une représentation, car ils obéissent aux mêmes relations de commutation; cela se vérifie en appliquant la conjugaison complexe :

$$([T_a, T_b])^* = -if_{abc} T_c^* \implies [-T_a^*, -T_b^*] = if_{abc} (-T_c^*) \quad (5.94)$$

Dans le cas de $SU(2)$, la représentation conjuguée est équivalente à la représentation d'origine, et le concept n'est pas utile. Ce n'est plus le cas dans $SU(3)$: la représentation conjuguée est généralement distincte, autrement dit, elle n'est pas reliée à la représentation d'origine par un simple changement de base. En particulier, le changement de signe fait que les valeurs propres de (T_3, T_8) associées à la représentation conjuguée sont opposées à celle de la représentation d'origine. Par exemple, la représentation conjuguée à la représentation fondamentale **3** de $SU(3)$ est notée **3^c**, et les valeurs propres associées sont

$$(m, y) = \left(-\frac{1}{2}, -\frac{1}{2\sqrt{3}}\right), \left(\frac{1}{2}, -\frac{1}{2\sqrt{3}}\right), \left(0, \frac{1}{\sqrt{3}}\right) \quad (5.95)$$

qui forment un triangle équilatéral droit dans le plan cartésien (figure 5.2).

Opérateurs d'échelle De même qu'on introduit opérateurs les d'échelle I_{\pm} dans l'étude des représentations de $SU(2)$, on peut introduire les opérateurs suivants en fonction

des générateurs de $SU(3)$:

$$I_{\pm} = T_1 \pm iT_2 \quad V_{\pm} = T_4 \pm iT_5 \quad U_{\pm} = T_6 \pm iT_7 \quad (5.96)$$

On calcule les relations de commutation suivantes avec T_3 et T_8 :

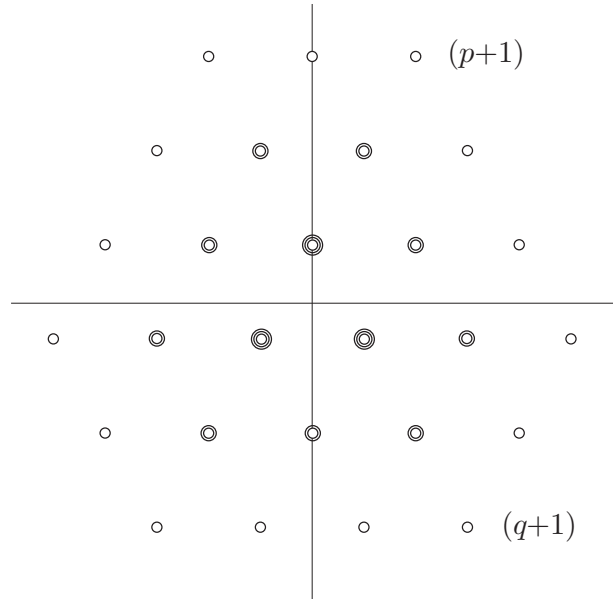
$$\begin{aligned} [T_3, I_{\pm}] &= \pm I_{\pm} & [T_8, I_{\pm}] &= 0 \\ [T_3, V_{\pm}] &= \pm \frac{1}{2} V_{\pm} & [T_8, V_{\pm}] &= \pm \frac{\sqrt{3}}{2} V_{\pm} \\ [T_3, U_{\pm}] &= \mp \frac{1}{2} U_{\pm} & [T_8, U_{\pm}] &= \pm \frac{\sqrt{3}}{2} U_{\pm} \end{aligned} \quad (5.97)$$

Ces relations confirment le caractère « échelle » de ces opérateurs. Autrement dit, l'opérateur I_+ augmente la valeur propre de T_3 de un, sans affecter la valeur propre de T_8 . Il produit donc un déplacement dans la direction $(1, 0)$ dans le plan (m, y) , alors que son conjugué I_- produit un déplacement dans la direction opposée. De même, V_{\pm} produit un déplacement dans la direction $\pm(1/2, \sqrt{3}/2)$ et U_{\pm} un déplacement dans la direction $\pm(-1/2, \sqrt{3}/2)$. Ces opérateurs d'échelle nous permettent donc de naviguer entre les différents états propres de chaque représentation.

Représentation adjointe La représentation dite *adjointe* de $SU(3)$, de dimension 8, est l'analogue pour ce groupe de la représentation d'isospin $I = 1$ pour le groupe $SU(2)$. Elle est en fait formée des différents déplacements possibles générés par les opérateurs d'échelle. Plus fondamentalement, le commutateur de deux générateurs T_a est une combinaison linéaire des générateurs. C'est le sens de l'expression $[T_a, T_b] = if_{abc} T_c$. Cette linéarité signifie qu'on pourrait représenter l'action d'un générateur T_a sur un autre générateur T_b par son commutateur avec lui, et former une matrice $(T_a)_{cb} = if_{abc}$. Ces matrices forment une représentation du groupe qu'on appelle *représentation adjointe*. Les états propres de cette représentation sont précisément les opérateurs d'échelle définis ci-haut, ainsi que les opérateurs de base T_3 et T_8 qui commutent entre eux, associés tous les deux à la valeur propre $(0, 0)$, doublement dégénérée. Voir la figure 5.2.

FIGURE 5.3

Représentation de dimension 42. Les valeurs propres représentées par les cercles simples sont non dégénérées. Celles représentées par les cercles doubles sont dégénérées deux fois, et celles représentées par les cercles triples sont dégénérées trois fois. Cette représentation est aussi notée $(2, 3)$.



Représentation générale Une représentation générale de $SU(3)$ a la forme d'un triangle tronqué dans le plan (m, y) , par exemple la représentation de dimension 42 illustrée à la figure 5.3. Si $p+1$ est le nombre de points sur le côté supérieur du diagramme et $q+1$ le nombre de points sur le côté inférieur, alors on peut désigner la représentation par le couple (p, q) . La figure 5.3, par exemple, illustre la représentation $(2, 3)$. La représentation fondamentale est $(1, 0)$, l'adjointe est $(1, 1)$. Les représentations dites *triangulaires* sont de la forme $(n, 0)$ où $n = 1, 2, 3, \dots$ et ont la forme d'un triangle équilatéral non tronqué. Leur dimension est $\frac{1}{2}n(n+1)$, c'est-à-dire 3, 6, 10, 15, 21, etc., et aucune des valeurs propres n'est dégénérée.

La représentation conjuguée à (p, q) est (q, p) . La dégénérescence des valeurs propres est déterminée par les règles suivantes :

1. Les valeurs propres situées sur la périphérie du diagramme sont non dégénérées.
2. La dégénérescence augmente de 1 à mesure qu'on pénètre dans le diagramme à partir de la périphérie, mais arrête d'augmenter dès qu'on rencontre une forme triangulaire à l'intérieur.

Les diagrammes des valeurs propres ont tous une symétrie par rotation de 120° , qui vient d'une symétrie fondamentale entre les trois opérateurs d'échelle I_+ , V_+ et U_+ .

5.C.3 Produits tensoriels de $SU(3)$

Lorsqu'on réunit deux systèmes physiques en un seul, l'espace des états en mécanique quantique est le produit tensoriel des espaces associés aux deux systèmes qu'on réunit. De même, si on dispose dans chacun de ces espaces d'une représentation d'un groupe de symétrie, les opérateurs du groupe agissant sur l'espace produit seront des produits tensoriels d'opérateurs :

$$R_1 = e^{-i\omega_a T_a^{(1)}} \quad \text{et} \quad R_2 = e^{-i\omega_a T_a^{(2)}} \quad \text{mènent à} \quad R_1 \otimes R_2 = e^{-i\omega_a T_a^{(\text{tot.})}} \quad (5.98)$$

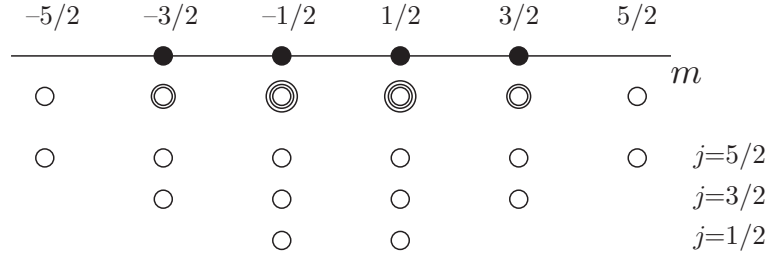
où on a défini le générateur total

$$T_a^{(\text{tot.})} = T_a^{(1)} \otimes 1 + 1 \otimes T_a^{(2)} \quad (5.99)$$

La théorie de l'addition des moments cinétiques est le prototype de la théorie plus générale du produit tensoriel des représentations.

Dans le cas de $SU(2)$, le produit tensoriel s'effectue simplement. Sans nous attarder à la question des coefficients de Clebsch-Gordan, étudions les valeurs propres de la somme des moments cinétiques. Supposons qu'on additionne deux moments cinétiques, dans les représentations j_1 et j_2 . Cette somme produira plusieurs représentations irréductibles de $SU(2)$, avec des valeurs de j allant de $j_1 + j_2$ à $|j_1 - j_2|$. On peut obtenir ce résultat en additionnant simplement les valeurs propres de T_3 pour les deux représentations, comme indiqué à la figure 5.4.

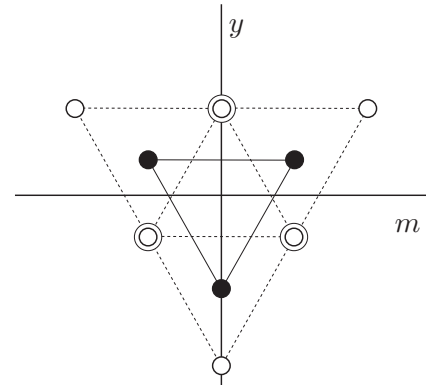
On procède de la même manière pour les produits tensoriels de $SU(3)$. Un exemple est illustré à la figure 5.5. On peut, en procédant à des produits tensoriels répétés, générer toutes les représentations irréductibles de $SU(3)$.


FIGURE 5.4

Addition des moments cinétiques, c'est-à-dire des générateurs du groupe, dans le cas de $SU(2)$. On combine ici une représentation de spin $j = \frac{3}{2}$ à une représentation de spin $j = 1$. On ajoute à chaque valeur propre de la première représentation les valeurs propres de la deuxième, ce qui produit un ensemble de $3 \times 4 = 12$ valeurs propres, la plupart dégénérées. On forme la représentation $j = \frac{5}{2}$ à partir de la plus élevée de ces valeurs propres. Si on enlève les valeurs propres correspondantes, il ne reste qu'une copie de la valeur propre $m = \frac{3}{2}$, à partir de laquelle on sait qu'une représentation $j = \frac{3}{2}$ existe dans la somme; si on enlève les valeurs propres correspondantes, il ne reste que les deux valeurs propres associées à la représentation $j = \frac{1}{2}$. Donc on sait, symboliquement, que $(\frac{3}{2}) \otimes (1) = (\frac{5}{2}) \oplus (\frac{3}{2}) \oplus (\frac{1}{2})$.

FIGURE 5.5

Combinaison de deux copies de la représentation fondamentale de $SU(3)$. On constate que $\mathbf{3} \otimes \mathbf{3} = \mathbf{6} + \mathbf{3}^c$.



On indique dans ce qui suit quelques produits tensoriels simples :

$$\begin{aligned}
 \mathbf{3} \otimes \mathbf{3} &= \mathbf{3}^c \oplus \mathbf{6} \\
 \mathbf{3} \otimes \mathbf{3}^c &= \mathbf{8} \oplus \mathbf{1} \\
 \mathbf{3} \otimes \mathbf{6} &= \mathbf{8} \oplus \mathbf{10} \\
 \mathbf{3}^c \otimes \mathbf{6} &= \mathbf{3} \oplus \mathbf{15} \\
 \mathbf{3} \otimes \mathbf{8} &= \mathbf{3} \oplus \mathbf{6}^c \oplus \mathbf{15} \\
 \mathbf{3} \otimes \mathbf{10} &= \mathbf{15} \oplus \mathbf{15}'
 \end{aligned} \tag{5.100}$$

où par définition la représentation $\mathbf{15}'$ est triangulaire et correspond à $(4, 0)$. Ces produits s'écriraient plutôt comme suit dans la notation (p, q) :

$$\begin{aligned}
 (1, 0) \otimes (1, 0) &= (0, 1) \oplus (2, 0) \\
 (1, 0) \otimes (0, 1) &= (1, 1) \oplus (0, 0) \\
 (1, 0) \otimes (2, 0) &= (1, 1) \oplus (3, 0) \\
 (0, 1) \otimes (2, 0) &= (1, 0) \oplus (2, 1) \\
 (1, 0) \otimes (1, 1) &= (1, 0) \oplus (0, 2) \oplus (2, 1) \\
 (1, 0) \otimes (3, 0) &= (2, 1) \oplus (4, 0)
 \end{aligned} \tag{5.101}$$

5.C.4 Symétrie de saveur des quarks

Retournons maintenant aux trois quarks u, d et s et au lagrangien (5.91). On considère ici ces trois quarks comme trois composantes d'un même objet (u, d, s) . Les interactions fortes traiteront ces trois quarks de la même manière, et si on néglige les différences de masses parmi ces trois quarks, alors la physique des interactions fortes devrait être identique pour les trois.

Les trois quarks (u, d, s) forment à eux trois la représentation fondamentale de $SU(3)$, soit $\mathbf{3} = (1, 0)$. Les trois antiquarks $(\bar{u}, \bar{d}, \bar{s})$, obtenus par conjugaison de charge, forment la représentation conjuguée $\mathbf{3}^c = (0, 1)$.

Les mésons sont obtenus en assemblant un quark et un antiquark. Les représentations correspondantes viennent du produit tensoriel de $\mathbf{3}$ et de $\mathbf{3}^c$, c'est-à-dire un octet $\mathbf{8} = (1, 1)$ et un singulet $\mathbf{1} = (0, 0)$. L'assignation des mésons à ces états est illustrée à la figure 5.6.

Les baryons sont obtenus en assemblant trois quarks. Les représentations correspondantes viennent du triple produit tensoriel de $\mathbf{3}$ avec lui-même, c'est-à-dire

$$\mathbf{3} \otimes \mathbf{3} \otimes \mathbf{3} = \mathbf{3} \otimes (\mathbf{3}^c \oplus \mathbf{6}) = \mathbf{1} \oplus \mathbf{8} \oplus \mathbf{8} \oplus \mathbf{10} \tag{5.102}$$

Le décuplet et l'octet sont illustrés à la figure 5.7.

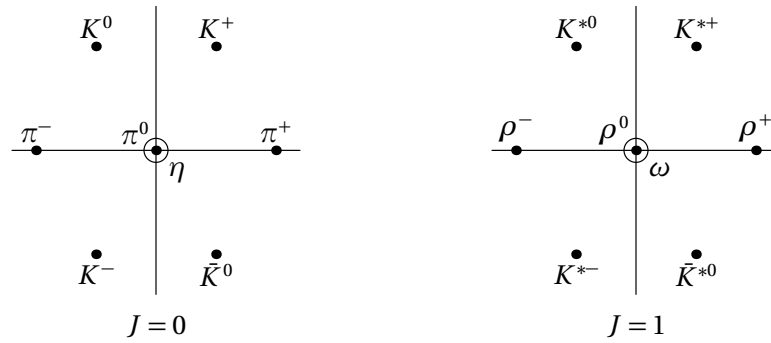


FIGURE 5.6

La classification des mésons dans des multiplets du $SU(3)$ de saveur. À gauche, les états de spin 0, à droite ceux de spin 1. Dans les deux cas, les octets sont illustrés. Il y a aussi un singulet η' de spin 0 et un singulet ϕ de spin 1.

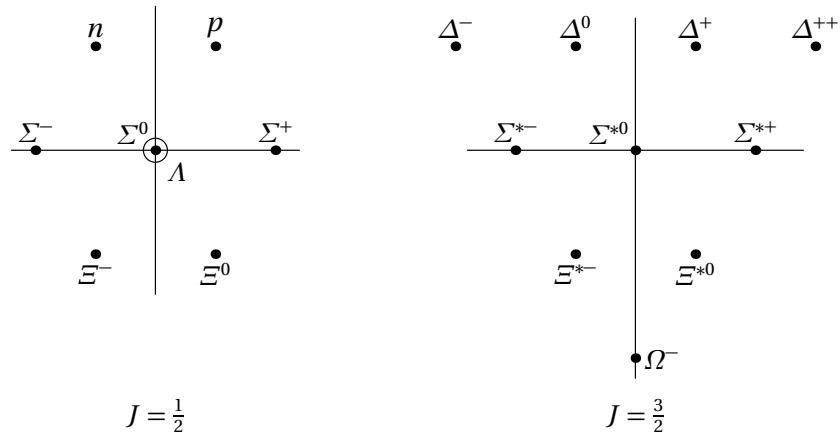


FIGURE 5.7

La classification des baryons dans des multiplets du $SU(3)$ de saveur. À gauche, les états de spin $\frac{1}{2}$ dans un octet; à droite ceux de spin $\frac{3}{2}$ dans un décuplet.

Relation avec les nombres quantiques B, Y, S, I_3 et Q À l'intérieur de chaque représentation, les propriétés comme le spin et la parité doivent être identiques. Nous avons donc ici une généralisation de l'isospin. Ce sont ces diagrammes, obtenus en premier à l'aide des nombres quantiques additifs que sont I_3 et l'étrangeté S , qui ont mené Gell-Mann à l'hypothèse des quarks. L'étrangeté est simplement l'opposé du nombre de quarks s . Les quarks u et d forment quant à eux un doublet d'isospin avec $I_3 = \pm \frac{1}{2}$. On définit aussi le nombre baryonique B , qui vaut un pour un baryon et zéro pour un méson. Ces nombres peuvent facilement être exprimés en fonction des nombres entiers N_u , N_d et N_s de chaque espèce de quarks :

$$B = \frac{1}{3}(N_u + N_d + N_s) \quad I_3 = \frac{1}{2}(N_u - N_d) \quad S = -N_s \quad (5.103)$$

D'où on tire, par inversion, que

$$N_u = \frac{1}{2}(3B + 2I_3 + S) \quad N_d = \frac{1}{2}(3B - 2I_3 + S) \quad (5.104)$$

et donc que la charge électrique (en multiples de e) est

$$Q = \frac{2}{3}N_u - \frac{1}{3}(N_d + N_s) = \frac{1}{2}(B + 2I_3 + S) \quad (5.105)$$

L'hypecharge est, quant à elle, donnée par :

$$Y = (S + B) \quad T_8 = \frac{\sqrt{3}}{2}Y \quad (5.106)$$

Les valeurs de S , B et I_3 pouvaient être inférées des observations, car ces nombres sont conservés par les interactions fortes et sont additifs.

La forte symétrie des diagrammes représentant les bayrons et les mésons a amené Gell-Mann à faire le lien avec les représentations du groupe $SU(3)$ et à supposer que ces représentations étaient les produits tensoriels de représentations plus simples (la représentation fondamentale et sa représentation conjuguée), ce qui sous-entend que les hadrons sont formés d'unité plus simples (les quarks), même si ces représentations fondamentales ne sont pas observées directement.

Formule de Gell-Mann & Okubo Comme la masse du quark s semble assez différente de celle des quarks u et d , la symétrie $SU(3)$ n'est qu'approximative et les différentes particules d'une même représentation n'ont pas la même masse. Cependant, la formule empirique suivante, dite de *Gell-Mann–Okubo*, représente les masses observées avec une précision de 1% :

$$M = M_0 + M_1 Y + M_2 \left(I(I+1) - \frac{1}{4} Y^2 \right) \quad (5.107)$$

où les paramètres $M_{0,1,2}$ peuvent être déterminés par lissage des masses de toutes les particules d'un multiplet donné. Cependant, cette formule n'a pas vraiment de pouvoir prédictif, sauf dans le cas du décuplet de baryons, car elle comporte trois paramètres ajustables ! Elle néglige bien sûr les différences de masses à l'intérieur de multiplets d'isospin, d'où son imprécision de 1%. La particule Ω^- , au bas du décuplet de baryons, a été prédite par Gell-Mann en 1962 sur la base de cette analyse. Elle fut découverte en 1964, avec les propriétés attendues (spin, parité, masse).

D Couleur et chromodynamique quantique

5.D.1 Symétrie de couleur

Peu après que Gell-Mann et G. ZWEIG aient proposé l'existence des quarks, et malgré le succès explicatif de cette hypothèse dans la classification des hadrons, certaines questions se posèrent :

1. Toutes les combinaisons possibles de quarks ne sont pas observées. On entend par là diverses combinaisons ayant toutes les valeurs de spin possible, etc. Dans le langage de la théorie des groupes – voir la section 5.C – le modèle des quarks prédit l'existence d'un décuplet de baryons, de deux octets et d'un singulet. Seul un des deux octets est observé. Pourquoi? Qu'est-ce qui fixe le spin de cet octet ($\frac{1}{2}$) et celui du décuplet ($\frac{3}{2}$)? Pourquoi n'observe-t-on pas un décuplet de spin $\frac{1}{2}$ et des octets de spin $\frac{3}{2}$?
2. La particule Δ^{++} , au sommet du décuplet de baryons, comporte trois quarks u et un spin $\frac{3}{2}$. On doit en outre présumer que les quarks y sont liés dans un état de fonction d'onde symétrique dans l'espace, comme c'est généralement le cas pour un état fondamental. Ceci signifie que l'état à trois quarks u est complètement symétrique lors de l'échange de deux quarks, en contradiction avec le principe de Pauli.

Une solution à ce dilemme fut proposée en 1964 par O.W. GREENBERG. Il émit l'hypothèse que les quarks possèdent aussi une *couleur*, c'est-à-dire que chaque saveur de quark existe dans des combinaisons linéaires de trois composantes ψ_i ($i = 1, 2, 3$) (nous utilisons des indices latins du milieu de l'alphabet pour désigner la couleur, comme l'indice d'isospin ou de saveur aux sections 5.B et 5.C; il ne faut cependant pas confondre les deux concepts). Autrement dit, une symétrie supplémentaire existe, par laquelle les trois variétés de chaque espèce de quarks sont combinées par une transformation unitaire. Contrairement à la symétrie d'isospin ou à la symétrie de saveur, cette symétrie est exacte et non approximative.

Ainsi, l'existence du Δ^{++} perd son mystère dès lors qu'on suppose qu'il est composé de trois quarks de couleurs différentes, et donc le principe de Pauli doit être appliqué en tenant compte de la couleur. Spécifiquement, on peut supposer que l'état quantique représentant le Δ^{++} est le produit tensoriel d'une partie orbitale, d'une partie spin et d'une partie « couleur ». Lors d'une permutation des trois quarks, la partie orbitale et la partie spin sont symétriques, alors que la partie couleur est antisymétrique, en accord avec le principe de Pauli. Cette dernière prend la forme suivante :

$$\frac{1}{\sqrt{6}}(|R\rangle \otimes |G\rangle \otimes |B\rangle + |G\rangle \otimes |B\rangle \otimes |R\rangle + |B\rangle \otimes |R\rangle \otimes |G\rangle - |G\rangle \otimes |R\rangle \otimes |B\rangle - |R\rangle \otimes |B\rangle \otimes |G\rangle - |B\rangle \otimes |G\rangle \otimes |R\rangle) \quad (5.108)$$

ou encore, sous une forme un peu abrégée,

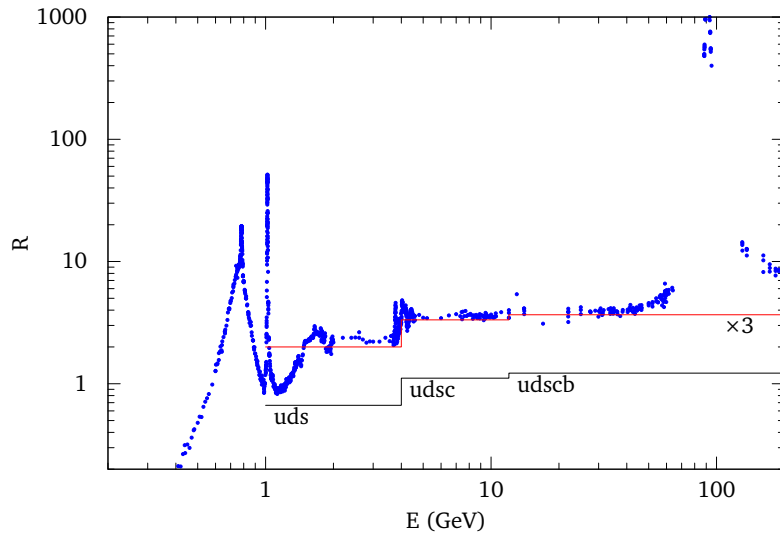
$$\frac{1}{\sqrt{6}}(|RGB\rangle + |GBR\rangle + |BRG\rangle - |GRB\rangle - |RBG\rangle - |BGR\rangle) = \frac{1}{\sqrt{6}}\epsilon_{ijk}|ijk\rangle \quad (5.109)$$

où les indices i, j, k prennent indépendamment les valeurs 1, 2 et 3 (ou R, G et B) et on somme, comme d'habitude, sur les indices répétés. On voit sans peine que cet état de couleur est complè-

tement antisymétrique lors d'une permutation des trois particules, ce qui permet de respecter le principe de Pauli.

FIGURE 5.8

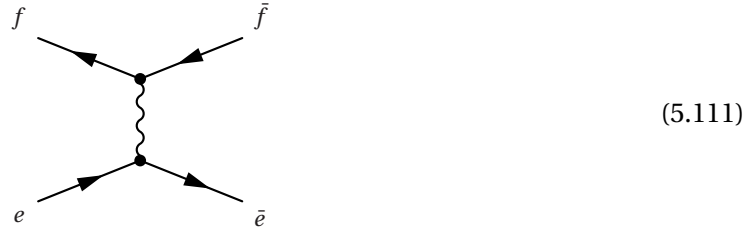
Rapport R entre les sections efficaces $\sigma_{e^+e^- \rightarrow \text{hadrons}}$ et $\sigma_{e^+e^- \rightarrow \mu^+\mu^-}$, en fonction de l'énergie E dans le repère du centre de masse. Notez les échelles logarithmiques sur les deux axes.



Le rapport R Une autre évidence expérimentale de l'existence de la couleur est le rapport R entre la section efficace totale des processus $e^+e^- \rightarrow \text{hadrons}$ et la section efficace du processus $e^+e^- \rightarrow \mu^+\mu^-$:

$$R = \frac{\sigma_{e^+e^- \rightarrow \text{hadrons}}}{\sigma_{e^+e^- \rightarrow \mu^+\mu^-}} \quad (5.110)$$

Chacun de ces processus peut être calculé à partir du diagramme suivant :



où f représente soit un muon, soit un quark, et \bar{f} est son antiparticule. Il faut bien sûr que l'énergie E de centre de masse de la paire $e\bar{e}$ soit suffisante pour créer la paire $f\bar{f}$. Le vertex supérieur est proportionnel à Q , la charge de la particule f . La section efficace s'obtient du carré de l'amplitude, de sorte que le rapport R ci-dessus devrait être égal à

$$R = \sum_a Q_a^2 \quad (5.112)$$

où Q_a est la charge du quark de type a ($a = u, d, s, c, b, t$) en multiple de la charge élémentaire. Ce résultat n'est valable que pourvu que la théorie des perturbations soit applicable, c'est-à-dire loin des résonnances qui correspondent à des états liés quark-antiquark. Seuls les quarks dont les masses sont suffisamment petites participent à cette somme, de sorte que le rapport R augmente avec l'énergie E .

La figure 5.8 montre la valeur mesurée de ce rapport en fonction de l'énergie de centre de masse de la paire $e\bar{e}$. L'escalier en noir représente la valeur de la somme (5.112), qui augmente au fur et à mesure qu'une nouvelle saveur de quark est disponible. On constate que cette somme est insuffisante pour bien représenter les données hors des résonances. L'escalier rouge est obtenu en multipliant la somme (5.112) par trois, pour tenir compte des trois couleurs de chaque saveur de quark. On constate que cette fois, l'accord avec les données est acceptable.

La neutralité électrique Un autre argument, plus théorique cette fois, pour l'existence de la couleur est la neutralité électrique de l'univers. Nous avons vu à la section 3.C.4 qu'à chaque espèce de fermion on peut associer une charge Q_0 , la charge contenue dans la mer de Dirac. Or la somme de ces charges sur toutes les espèces de fermions devrait s'annuler, par neutralité électrique. C'est précisément ce qui se produit, si on tient compte de la couleur : la somme des charges électriques des quarks dans une famille donnée est alors $3 \times (\frac{2}{3} - \frac{1}{3}) = 1$, ce qui compense la charge -1 de l'électron.

États neutres en couleur Non seulement l'état (5.109) est-il complètement antisymétrique, mais il est aussi invariant lors d'une transformation de couleur, c'est-à-dire d'une rotation unitaire dans l'espace des couleurs. Lors d'une telle transformation, chaque quark subit une rotation de couleur par une matrice unitaire U :

$$|i\rangle \rightarrow U_{ij}|j\rangle \quad (5.113)$$

L'état à trois quarks subit la même rotation pour chacun des trois quarks, ce qui se traduit par

$$|\psi\rangle \rightarrow |\psi'\rangle = \frac{1}{\sqrt{6}} \varepsilon_{ijk} U_{il} U_{jm} U_{kn} |lmn\rangle \quad (5.114)$$

Or, il est facile de vérifier que, pour une matrice quelconque M ,

$$\varepsilon_{ijk} M_{il} M_{jm} M_{kn} = \varepsilon_{lmn} \det M \quad (5.115)$$

On trouve donc

$$|\psi'\rangle = \varepsilon_{lmn} \det U |lmn\rangle = \det U |\psi\rangle \quad (5.116)$$

Comme la matrice U est unitaire, son déterminant est un nombre unimodulaire ($|\det U| = 1$) et donc $|\psi'\rangle$ est égal à $|\psi\rangle$, à une phase près. De plus, si on impose une contrainte supplémentaire à la matrice U , à savoir que son déterminant est l'unité ($\det U = 1$), alors l'état est vraiment invariant lors de cette transformation de couleur. L'ensemble des matrices unitaires d'ordre 3 et de déterminant unité forme un *groupe* appelé $SU(3)$ (voir l'annexe 7.D ou la section 5.C).

Bien que le mot « couleur » soit ici une métaphore, on la pousse un peu plus loin en parlant d'états « blancs » pour désigner les états neutres en couleurs, c'est-à-dire invariants lors d'une rotation dans l'espace des couleurs. Existe-t-il d'autres états neutres, mis à part l'état ci-dessus? On peut montrer, grâce aux techniques de la théorie des groupes, que :

1. l'état (5.109) est le seul état neutre en couleurs composé de trois quarks.
2. Il n'y a pas d'états neutres en couleurs composés de deux, quatre, ou cinq quarks.
3. Des états neutres en couleurs sont évidemment possibles pour un multiple de trois quarks.
4. Il existe un état neutre en couleur pour un objet composé d'un quark et d'un antiquark.

Élaborons sur ce dernier point.

État quark-antiquark Lors d'une rotation dans l'espace des couleurs, un état formé d'un antiquark se transforme ainsi : $|\bar{i}\rangle \rightarrow U_{ik}^* |\bar{k}\rangle$, où la barre sur l'indice (\bar{i}) est utilisée pour insister sur le fait qu'il s'agit d'un antiquark. L'utilisation du conjugué complexe de la matrice U vient du fait que le développement en modes du champ de Dirac ψ fait intervenir les opérateurs d'annihilation des particules, et de création des antiparticules, de sorte que, schématiquement,

$$c_{\mathbf{k}} \rightarrow U c_{\mathbf{k}} \quad \text{et} \quad d_{\mathbf{k}}^\dagger \rightarrow U d_{\mathbf{k}}^\dagger \quad (5.117)$$

et donc que $c_{\mathbf{k}}^\dagger |0\rangle$ se transforme avec la matrice conjuguée de celle qui transforme $d_{\mathbf{k}}^\dagger |0\rangle$. Sachant cela, on peut écrire l'état suivant pour un méson (c'est-à-dire un ensemble quark-antiquark) :

$$|\psi\rangle = \frac{1}{\sqrt{3}} (|R\bar{R}\rangle + |G\bar{G}\rangle + |B\bar{B}\rangle) \quad (5.118)$$

ou encore

$$|\psi\rangle = \frac{1}{\sqrt{3}} \sum_{i=1}^3 |i, \bar{i}\rangle \quad (5.119)$$

La somme sur i sera implicite ci-dessous, comme toute somme sur un indice répété. Lors d'une transformation de couleur,

$$|i\rangle \rightarrow U_{ij} |j\rangle \quad \text{et} \quad |\bar{i}\rangle \rightarrow U_{ik}^* |\bar{k}\rangle \quad (5.120)$$

et donc

$$|\psi\rangle \rightarrow |\psi'\rangle = U_{ij} U_{ik}^* |j, \bar{k}\rangle \quad (5.121)$$

Comme la matrice U est unitaire,

$$U_{ij} U_{ik}^* = (U^\dagger)_{ki} U_{ij} = (U^\dagger U)_{kj} = \delta_{kj} \quad (5.122)$$

et enfin

$$|\psi'\rangle = \delta_{kj} |j, \bar{k}\rangle = |j, \bar{j}\rangle = |\psi\rangle \quad (5.123)$$

Cet état est donc aussi invariant par une transformation de couleur.

Confinement de la couleur L'hypothèse de l'existence de la couleur prend un sens nouveau quand on lui ajoute la condition que seuls les états neutres sont observables. Ainsi, les combinaisons de quarks et/ou d'antiquarks qui sont absentes de la classification élaborée par Gell-Mann le seraient parce qu'elles ne peuvent pas être réalisées d'une manière neutre en couleur.

Se pose alors la question fondamentale suivante : d'où vient ce principe de la neutralité de couleur ? Son origine doit nécessairement être dynamique, c'est-à-dire qu'il doit découler d'une théorie fondamentale des interactions fortes, théorie encore inconnue en 1964. Han et Nambu proposèrent qu'une théorie de l'interaction forte puisse justement être basée sur cette symétrie $SU(3)_c$, tâche qui fut accomplie plus en détail par Gross, Wilcek et Politzer en 1973. Cette théorie est basée sur l'idée d'une symétrie de jauge, généralisation de l'électrodynamique quantique, proposée par YANG et MILLS en 1954, que nous allons maintenant expliquer.

5.D.2 Théorie de Yang-Mills

En 1954, Yang et Mills ont généralisé l'idée d'une symétrie de phase locale à une symétrie plus vaste impliquant des rotations unitaires de plusieurs champs. La théorie obtenue est appelée *théorie de jauge non abélienne*, car elle implique des transformations qui ne commutent pas entre elles (des transformations unitaires, par exemple) et que ces transformations forment un groupe non commutatif, ou groupe *non abélien*. Ces théories forment maintenant la base du modèle standard des interactions fondamentales.

Nous adopterons comme modèle la théorie dynamique de la couleur, ou chromodynamique, qui est la théorie fondamentale des interactions fortes entre quarks. Considérons donc un spineur ψ décrivant un quark dans ses trois couleurs, c'est-à-dire que le spineur $\psi_{\alpha i}$ comporte un indice de Dirac ($\alpha = 1, 2, 3, 4$) et un indice de couleur ($i = 1, 2, 3$). L'indice de Dirac sera toujours implicite, alors que l'indice de couleur sera implicite la plupart du temps. Les transformations de jauge agissant sur ce spineur sont en fait des rotations unitaires dans l'espace des couleurs, autrement dit

$$\psi_i \rightarrow \psi'_i = U_{ij} \psi_j \quad \text{ou} \quad \psi \rightarrow \psi' = U \psi \quad (5.124)$$

où U est une matrice unitaire.

Le lagrangien de Dirac

$$\mathcal{L} = \bar{\psi}(i\gamma^\mu \partial_\mu - m)\psi \quad (5.125)$$

est certainement invariant lors d'une transformation de couleur globale. Autrement dit, si U est une matrice unitaire agissant dans l'espace des couleurs et indépendante de la position et du temps, alors le lagrangien est invariant lorsqu'on procède à la transformation $\psi \rightarrow \psi' = U\psi$, car alors $\bar{\psi}' = \bar{\psi}U^\dagger$ et les deux facteurs U^\dagger et U se combinent pour former la matrice unité. Notons bien sûr que la matrice U agissant dans l'espace des couleurs seulement, commute avec toutes les matrices de Dirac, qui agissent dans l'espace des spineurs seulement.

Transformation de jauge Cependant, le lagrangien n'est plus invariant si la matrice U dépend de la position ou du temps, en raison de la dérivée agissant sur ψ . Nous devons définir encore une fois une dérivée covariante $\mathcal{D}_\mu = \partial_\mu + igA_\mu$, où g est une constante de couplage (l'analogue dans les interactions fortes de la charge électrique). Il faut définir la transformation de jauge du champ A_μ de manière à ce que la dérivée covariante se transforme simplement, comme le champ ψ lui-même :

$$\mathcal{D}_\mu \psi \rightarrow \mathcal{D}'_\mu \psi' = U(\mathcal{D}_\mu \psi) \quad (5.126)$$

Dans le détail, ceci entraîne que

$$(\partial_\mu + igA'_\mu)U\psi = (\partial_\mu U)\psi + U\partial_\mu \psi + igA'_\mu U\psi \quad (5.127)$$

doit être égal à

$$U(\partial_\mu + igA_\mu)\psi \quad (5.128)$$

et donc que

$$\partial_\mu U + igA'_\mu U = igUA_\mu \implies \boxed{A'_\mu = UA_\mu U^\dagger + \frac{i}{g} \partial_\mu U U^\dagger} \quad (5.129)$$

Le champ A_μ est appelé *champ de jauge* et est manifestement une matrice du même ordre que U .

L'action de Dirac invariante de jauge est alors simplement obtenue par couplage minimal, c'est-à-dire en remplaçant les dérivées ordinaires par des dérivées covariantes :

$$\mathcal{L} = \bar{\psi}(i\gamma^\mu \mathcal{D}_\mu - m)\psi = \bar{\psi}(i\gamma^\mu \partial_\mu - g A_\mu \gamma^\mu - m)\psi \quad (5.130)$$

5.D.3 Champ de jauge et générateurs

L'objectif de cette section est de montrer que le champ de jauge A_μ figurant dans l'éq. (5.129) doit être une matrice hermitienne et qu'il peut s'exprimer comme une combinaison linéaire des générateurs T_a du groupe $SU(3)$ (voir l'annexe 7.D) :

$$A_\mu = A_\mu^a T_a \quad (5.131)$$

Nous allons maintenant montrer que l'expression $i\partial_\mu U U^\dagger$ figurant dans l'éq. (5.129) est une combinaison linéaire réelle des générateurs T_a . Considérons une matrice U appartenant à $SU(n)$, ainsi qu'une légère variation de cette matrice, $U + \delta U$. Cette nouvelle matrice peut être exprimée comme le produit d'un élément du groupe très proche de l'identité, post-multiplié par U :

$$U + \delta U = (1 + i\delta\omega_a T_a)U \quad \text{et donc} \quad \delta U = i\delta\omega_a T_a U \quad \text{ou} \quad \delta U U^\dagger = i\delta\omega_a T_a \quad (5.132)$$

où les $\delta\omega_a$ sont des paramètres infinitésimaux. En supposant que la variation δU résulte d'un changement de position lors d'une transformation de jauge, on trouve simplement

$$\partial_\mu U U^\dagger = i\partial_\mu \omega_a T_a \quad \text{ou} \quad \frac{i}{g} \partial_\mu U U^\dagger = -\frac{1}{g} \partial_\mu \omega_a T_a \quad (5.133)$$

et ce terme de la relation (5.129) est donc une combinaison linéaire des générateurs. On en conclut que le champ de jauge A_μ est lui-même une combinaison $A_\mu = A_\mu^a T_a$ des générateurs, où les champs A_μ^a sont maintenant des champs réels non matriciels.

Cette conclusion est compatible avec le premier terme de la relation (5.129). En effet, en posant $H = -i\omega_a T_a$ et d'après la relation de Hausdorff, ce premier terme est

$$e^{-H} A_\mu e^H = A_\mu + [A_\mu, H] + \frac{1}{2!} [[A_\mu, H], H] + \frac{1}{3!} [[[A_\mu, H], H], H] + \dots \quad (5.134)$$

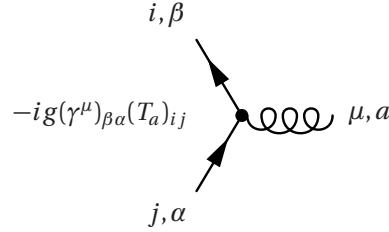
Comme H est une combinaison des générateurs T_a et que le commutateur de deux combinaisons linéaires des T_a est encore une combinaison linéaire des T_a ,² toute l'expression ci-haut est encore une combinaison linéaire des générateurs.

2. Le commutateur de deux matrices hermitiennes est une matrice anti-hermitienne, soit i multiplié par une matrice hermitienne.

Vertex de la QCD Le terme du lagrangien qui couple le champ de jauge aux quarks prend alors la forme

$$\mathcal{L}_I = -g \bar{\psi} A_\mu \gamma^\mu \psi = -g A_\mu^a \bar{\psi} \gamma^\mu T_a \psi = -g A_\mu^a \bar{\psi}_{ai} (\gamma^\mu)_{\alpha\beta} (T_a)_{ij} \psi_{\beta j} \quad (5.135)$$

où, dans la dernière expression, nous avons explicité les indices de Dirac (α, β) et les indices de couleur (i, j) qui étaient précédemment implicites. La règle de Feynman associée à cette interaction serait d'inclure un vertex accompagné des facteurs suivants :



$$-ig(\gamma^\mu)_{\beta\alpha}(T_a)_{ij} \quad (5.136)$$

Notez que le champ de jauge (appelé *gluon* dans la QCD) est représenté par une ligne en forme de ressort, pour le distinguer du photon et pour accentuer le caractère plus « fort » de l'interaction. Le gluon porte aussi de la couleur, mais en combinaisons de « couleurs » et d'« anticouleurs ». Il y a 9 combinaisons possibles, qui ne sont cependant pas linéairement indépendantes; les huit combinaisons linéairement indépendantes sont justement décrites par les matrices T_a .

5.D.4 Le tenseur de Faraday

Une fois la symétrie de jauge accomplie pour le champ de Dirac et le champ de jauge A_μ , il faut construire une théorie dynamique du champ A_μ lui-même, modelée sur la théorie électromagnétique. Plus précisément, il faut construire une densité lagrangienne pour A_μ , en suivant autant que possible le modèle électromagnétique et en s'assurant que cette densité lagrangienne soit également invariante de jauge.

La première étape consiste à construire le tenseur de Faraday

$$F_{\mu\nu} = \partial_\mu A_\nu - \partial_\nu A_\mu \quad (5.137)$$

Malheureusement, cette expression n'est pas invariante de jauge dans le cas non abélien. La solution est de remplacer les dérivées ordinaires par des dérivées covariantes :

$$\begin{aligned} F_{\mu\nu} &= \mathcal{D}_\mu A_\nu - \mathcal{D}_\nu A_\mu \\ &= \partial_\mu A_\nu - \partial_\nu A_\mu + ig[A_\mu, A_\nu] \end{aligned} \quad (5.138)$$

Voyons comment cette expression se comporte lors d'une transformation de jauge :

$$F_{\mu\nu} \rightarrow F'_{\mu\nu} = \mathcal{D}'_\mu A'_\nu - \mathcal{D}'_\nu A'_\mu \quad (5.139)$$

Calculons d'abord

$$\mathcal{D}'_\mu A'_\nu = (\partial_\mu + igUA_\mu U^\dagger - \partial_\mu UU^\dagger) \left(UA_\nu U^\dagger + \frac{i}{g} \partial_\nu UU^\dagger \right)$$

$$\begin{aligned}
 &= \partial_\mu U A_\nu U^\dagger + U \partial_\mu A_\nu U^\dagger + U A_\nu \partial_\mu U^\dagger + \frac{i}{g} \partial_\mu \partial_\nu U U^\dagger \\
 &\quad + \frac{i}{g} \partial_\nu U \partial_\mu U^\dagger + i g U A_\mu A_\nu U^\dagger + U A_\mu \partial_\nu U^\dagger - \partial_\mu U A_\nu U^\dagger + \frac{i}{g} \partial_\mu U \partial_\nu U^\dagger
 \end{aligned} \tag{5.140}$$

où nous avons utilisé à deux reprises le fait que $U^\dagger \partial_\mu U = \partial_\mu (U^\dagger U) - \partial_\mu U^\dagger U = -\partial_\mu U^\dagger U$. Notons que les premier et huitième termes se compensent mutuellement. D'autre part, comme nous devons soustraire de cette expression une expression équivalente où les indices μ et ν sont échangés, nous pouvons éliminer de cette expression tous les termes qui sont symétriques lors de l'échange $\mu \leftrightarrow \nu$. Or, le troisième terme se combine avec le septième pour former une expression symétrique en (μ, ν) . La même chose est vraie des cinquième et neuvième termes. Enfin, le quatrième terme est manifestement symétrique de lui-même. Il reste les deuxième et sixième termes :

$$U \partial_\mu A_\nu U^\dagger + i g U A_\mu A_\nu U^\dagger \tag{5.141}$$

En échangeant $\mu \leftrightarrow \nu$ et en soustrayant, on trouve

$$F'_{\mu\nu} = U(\partial_\mu A_\nu - \partial_\nu A_\mu)U^\dagger + i g U[A_\mu, A_\nu]U^\dagger = U F_{\mu\nu} U^\dagger \tag{5.142}$$

Donc le tenseur de Faraday $F_{\mu\nu}$ n'est pas invariant de jauge, mais se transforme de manière simple.

En fonction des générateurs T_a , le tenseur de Faraday s'écrit

$$F_{\mu\nu} = (\partial_\mu A_\nu^a - \partial_\nu A_\mu^a) T_a + i g A_\mu^b A_\nu^c [T_b, T_c] \tag{5.143}$$

Comme $[T_b, T_c] = i f_{bca} T_a$, on peut écrire

$$F_{\mu\nu} = F_{\mu\nu}^a T_a \quad \text{où} \quad \boxed{F_{\mu\nu}^a = \partial_\mu A_\nu^a - \partial_\nu A_\mu^a - g f_{bca} A_\mu^b A_\nu^c} \tag{5.144}$$

5.D.5 Le lagrangien de la QCD

Le lagrangien du champ électromagnétique est

$$\mathcal{L}_{EM} = -\frac{1}{4} F_{\mu\nu} F^{\mu\nu} \tag{5.145}$$

Cette forme doit être adaptée au cas non abélien, car en soi elle n'est pas invariante de jauge (en fait, ce n'est pas un scalaire, mais une matrice dans ce cas). La solution consiste à prendre la trace :

$$\mathcal{L}_{QCD} = -\frac{1}{2} \text{tr}(F_{\mu\nu} F^{\mu\nu}) \tag{5.146}$$

Ce lagrangien est bel et bien invariant de jauge, car

$$\mathcal{L}'_{QCD} = -\frac{1}{2} \text{tr}(F'_{\mu\nu} F'^{\mu\nu}) = -\frac{1}{2} \text{tr}(U^\dagger F_{\mu\nu} U U^\dagger F^{\mu\nu} U) = -\frac{1}{2} \text{tr}(F_{\mu\nu} F^{\mu\nu}) \tag{5.147}$$

en raison de la propriété cyclique de la trace d'un produit de matrices : $\text{tr}(ABC) = \text{tr}(BCA) = \text{tr}(CAB)$.

En fonction des composantes de couleur $F_{\mu\nu}^a$ du tenseur de Faraday, le lagrangien s'exprime ainsi :

$$\mathcal{L}_{QCD} = -\frac{1}{2} F_{\mu\nu}^a F^{b\mu\nu} \text{tr}(T_a T_b) \quad (5.148)$$

La convention habituelle sur la normalisation des générateurs est de les choisir tels que

$$\text{tr}(T_a T_b) = \frac{1}{2} \delta_{ab} \quad (5.149)$$

On trouve alors

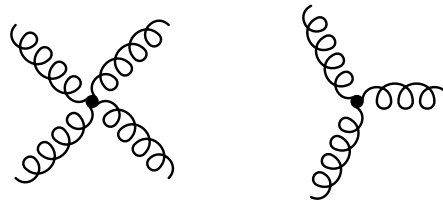
$$\mathcal{L}_{QCD} = -\frac{1}{4} F_{\mu\nu}^a F^{a\mu\nu} \quad (5.150)$$

Dans la limite où la constante de couplage g tend vers zéro, $F_{\mu\nu}^a$ devient simplement équivalent au tenseur de Faraday électromagnétique, et alors on se trouve en présence d'une théorie équivalente à N copies de l'électromagnétisme, N étant le nombre de générateurs (8 pour la QCD). Ainsi, N bosons, similaires au photon, apparaissent dans la théorie. En QCD, ces bosons sont appelés *gluons*, et ils transmettent l'interaction forte de la même manière que les photons transmettent l'interaction électromagnétique. Par exemple, les quarks pourraient diffuser l'un sur l'autre par l'échange d'un gluon virtuel, etc.

Si g est non nul, cependant, les choses sont plus complexes. Le lagrangien comporte des termes cubiques et des termes quartiques en A , comme le montre l'expression explicite suivante :

$$\mathcal{L}_{QCD} = -\frac{1}{4} \left(\partial_\mu A_\nu^a - \partial_\nu A_\mu^a - g f_{bca} A_\mu^b A_\nu^c \right) \left(\partial^\mu A^{a\nu} - \partial^\nu A^{a\mu} - g f_{dea} A^{d\mu} A^{e\nu} \right) \quad (5.151)$$

Les termes cubiques en A sont d'ordre g , et le terme quartique d'ordre g^2 . Ce lagrangien, ou le hamiltonien correspondant, est impossible à résoudre. Autrement dit, les états propres de cette théorie ne sont pas connus, et en tout cas ne correspondent pas à des bosons libres, comme dans le cas du champ électromagnétique. Bref, même en l'absence de quarks, la QCD est une théorie complexe. Si g était petit, alors on pourrait utiliser la théorie des perturbations (les diagrammes de Feynman) pour décrire les effets de termes cubiques et quartiques. Des vertex comme ceux illustrés ci-dessous seraient des ingrédients des diagrammes de Feynman de la QCD.


(5.152)

Le lagrangien complet de la QCD comporte l'expression ci-haut (5.151), plus les termes décrivant les quarks et leur interaction avec les gluons, soit l'expression (5.130).

5.D.6 Propriétés de la QCD

Terminons ce chapitre avec quelques remarques sur la QCD et les théories de jauge en général.

Une propriété de la QCD, commune à toutes les théories de jauge, est la renormalisabilité. Une théorie quantique des champs, pour être utile et bien définie, doit être renormalisable (voir la discussion de la p. 124). Il se trouve que toutes les théories de jauge sont renormalisables.

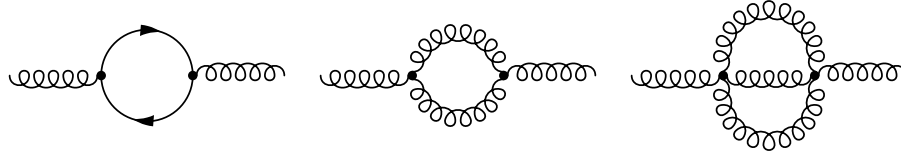


FIGURE 5.9

Diagrammes contribuant à la polarisation de couleur du vide en QCD. Celui de gauche représente l'émission de paires virtuelles quark-antiquark et les deux autres l'émission de gluons virtuels. Les deux contributions ($q\bar{q}$ et gluons) ont des signes opposés.

Liberté asymptotique L'un des effets de cette procédure de renormalisation est que la constante de couplage g n'est pas vraiment une constante. Elle dépend faiblement d'une échelle d'énergie q , échelle d'énergie caractéristique des phénomènes étudiés. Il se trouve que la constante $\alpha_s = g^2/4\pi$ se comporte comme suit en fonction de l'échelle d'énergie q :

$$\frac{1}{\alpha_s(q^2)} = \frac{1}{\alpha_s(q_0^2)} + \frac{33 - 2N_q}{12\pi} \ln \frac{q^2}{q_0^2} \quad (5.153)$$

où N_q est le nombre de saveurs de quarks (soit 6 dans le modèle standard). Ce résultat est obtenu par la théorie des perturbations (diagrammes de Feynman à une boucle) et n'est valable que si α_s est petit, mais indique que la constante de couplage tend vers zéro aux très grandes énergies. Cette propriété porte le nom de *liberté asymptotique*. Les diagrammes de Feynman qui mènent à cette formule sont illustrés à la figure 5.9. Le « 33 » provient de la contribution des boucles de gluons, alors que le $-2N_q$ provient des boucles de quarks et possède un signe opposé (les règles de Feynman imposent un facteur -1 aux boucles de fermions). Si les gluons n'interagissaient pas entre eux, le coefficient du logarithme serait négatif, ce qui entraînerait que la constante de couplage augmente logarithmiquement avec l'énergie. C'est effectivement ce qui se produit en électrodynamique (QED). Comme le nombre de familles de quarks est $N_q = 6$, le préfacteur $33 - 2N_q$ est positif et cela mène à la liberté asymptotique. L'une des conséquences de la liberté asymptotique est que lors de processus impliquant un très grand transfert de quantité de mouvement, comme dans les expériences de diffusion inélastique profonde, la constante de couplage effective est petite et les quarks se comportent comme des particules quasi libres.

Selon l'expression (5.153), il existe une échelle d'énergie Λ à laquelle la constante de couplage devient infinie. On peut alors écrire, de manière équivalente, pour $q^2 > \Lambda^2$,

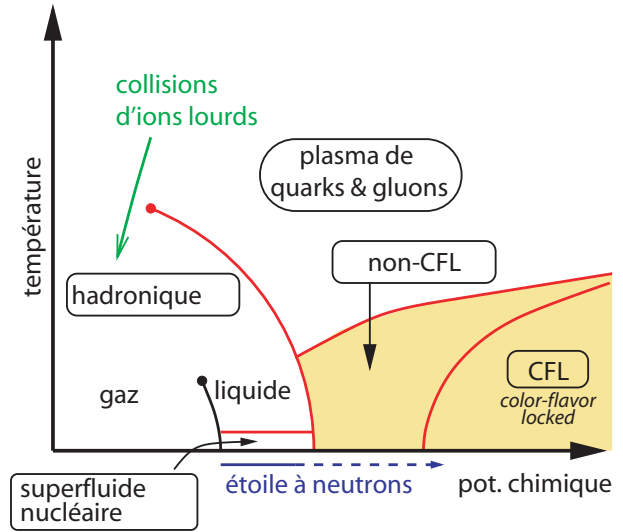
$$\alpha_s = \frac{12\pi}{(33 - 2N_q) \ln(q^2/\Lambda^2)} \quad (5.154)$$

On estime Λ à 217 ± 24 MeV (la précision n'est pas énorme).

Une façon de présenter la liberté asymptotique est de parler d'effet anti-écran (voir la discussion de l'effet d'écrantage en page 123). En QCD, la polarisation de couleur du vide par émission de paires

FIGURE 5.10

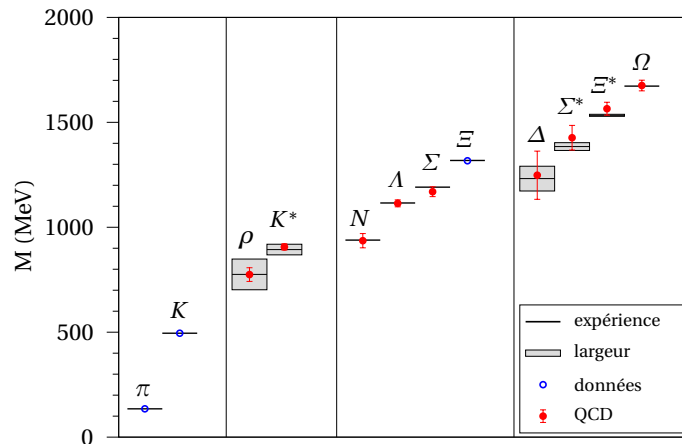
Diagramme de phase supposé de la matière hadronique, en fonction du potentiel chimique μ (relié à la densité) et de la température. Voir M.G. Alford et al., Rev. Mod. Phys. **80**, 1455 (2008).



virtuelles quark-antiquark agit dans le même sens qu'en électrodynamique, et donc contre la liberté asymptotique. Par contre, une polarisation de couleur dans le sens opposé (donc un anti-écranage) est causée par les gluons (les deux diagrammes du bas de la figure 5.9), et la contribution des gluons domine. Au total, l'effet anti-écran domine et il y a liberté asymptotique.

FIGURE 5.11

Masses relatives des hadrons légers calculées à l'aide de la QCD sur réseau. Les points rouges sont les prédictions de la QCD, les points bleus sont des observations, pour certaines particules, qui servent de base aux prédictions pour d'autres particules. La masse du proton et du neutron (N) est prédite à 2% près (la différence de masse entre les deux particules n'est pas résolue dans ce calcul). Données tirées de Dürr *et al.*, Science, **322** 1224 (2008).



QCD sur réseau La contrepartie de la liberté asymptotique est l'augmentation présumée de g^2 quand l'échelle d'énergie diminue. On peut présumer que, en deçà d'une certaine échelle d'énergie Λ , la théorie des perturbations perd toute son utilité et qu'il est vain de penser en termes de diagrammes de Feynman. D'autres approches théoriques, dites *non perturbatives*, sont requises. De telles méthodes, en particulier numériques, sont un sujet de recherche très actif en physique théorique des particules. L'une de ces approches, appelée *QCD sur réseau* (angl. *lattice QCD*), consiste à résoudre les équations de la QCD sur un maillage fini de points dans l'espace temps (le réseau). Le formalisme de *l'intégrale de chemins* de la mécanique quantique est utilisé en conjonction avec la méthode d'échantillonnage Monte-Carlo. Cette méthode requiert une quantité importante de ressources informatiques et elle fait face à de nombreux défis techniques. Cependant, elle a connu des succès importants depuis quelques années. En particulier, les masses relatives des hadrons légers ont pu être calculées *ab initio* avec une précision de l'ordre de 2%. Voir

la figure 5.11.

Confinement de la couleur Le revers de la médaille de la liberté asymptotique est le *confinement* de la couleur : on s'attend à ce que l'interaction forte, dans son régime non perturbatif, interdise les états de basse énergie qui ne sont pas neutres en couleur. En particulier, le gluon n'est pas neutre en couleur : il est affecté par une transformation de couleur, comme on peut le constater par la forme de la transformation de jauge du champ A_μ^a (eq. (5.129)), qui est non triviale même quand la matrice U est constante. Donc, le gluon, en tant que particule, ne peut exister à l'état libre.

Boules de glu Par contre, il est théoriquement possible que des états composés de deux ou trois gluons puissent exister. Ces particules composites, appelées «boules de glu» (angl. *glueballs*), auraient des masses s'échelonnant de 1,4 GeV à 5 GeV selon les prédictions de la QCD sur réseau. Par contre, elles seraient instables, se désintégrant très rapidement vers des mésons légers comme des pions. Ceci rend leur détection difficile, de sorte qu'à ce jour aucune confirmation solide de leur existence n'a été rapportée.

Diagramme de phase de la QCD Les propriétés présumées de la QCD, provenant d'un éventail de méthodes théoriques, sont résumées dans le *diagramme de phase* illustré sur la figure 5.10. Ce diagramme distingue différentes phases de la matière hadronique en fonction du potentiel chimique μ et de la température T . Le potentiel chimique est une fonction monotone de la densité. Un parallèle existe entre une partie de ce diagramme de phase et celui d'une substance pure, comme l'eau par exemple : il comporte des changements de phase du premier ordre (comme la passage d'un liquide à un gaz) qui se terminent par des points critiques. La phase située en haut du diagramme est un plasma de quarks et gluons ; c'est la phase à haute température qui prévalait dans les premières microsecondes de l'Univers et qui est reproduite de manière furtive dans les collisions d'ions lourds : les quarks et gluons se déplacent quasi librement, c'est-à-dire que la couleur n'est pas confinée. Dans le coin inférieur gauche du diagramme se trouve un gaz formé de « molécules » : les hadrons (baryons et mésons). La couleur est *confinée* dans cette phase : seuls les états neutres de couleurs peuvent exister. En augmentant la densité, toujours à basse température, on transite vers un liquide, soit la matière nucléaire elle-même (un ensemble de baryons). Les noyaux réels sont en quelque sorte des gouttelettes de cette phase, mais l'intérieur d'une étoile à neutrons constitue un très grand volume de cette phase. Éventuellement, en poussant la densité encore plus loin (peut-être comme au cœur des étoiles à neutrons), on atteint un régime où les interactions sont effectivement plus faibles et où on peut appliquer certaines méthodes de la physique de la matière condensée ; ces méthodes prédisent l'apparition d'une phase supraconductrice de couleur (CFL) où des paires de Cooper de quarks se forment et condensent. Cette phase n'a assurément pas encore donné signe de vie, mais son existence est une conséquence pratiquement inévitable de la théorie.

E Problèmes

Problème 5.1 Parité

On définit l'opération de parité – ou inversion de l'espace – par la transformation $\mathbf{r} \rightarrow -\mathbf{r}$. L'effet de cette transformation sur le champ de Dirac est le suivant :

$$\psi(\mathbf{r}, t) \rightarrow \psi'(\mathbf{r}, t) = \eta \gamma^0 \psi(-\mathbf{r}, t) \quad (5.155)$$

où $\eta = \pm 1$ est la *parité intrinsèque* du champ de Dirac.

Partez de l'action associée au champ de Dirac :

$$S = \int d^4x (i \bar{\psi} \gamma^\mu \partial_\mu \psi - m \bar{\psi} \psi) \quad (5.156)$$

et démontrez que cette action est invariante lorsqu'on procède à une transformation de parité.

Problème 5.2 Matrice γ^5

Démontrez les propriétés (5.9).

Problème 5.3 Spineurs chiraux

Vérifiez l'équation (5.14) en substituant les expressions des projections droite et gauche et en utilisant les propriétés de la matrice γ^5 .

Problème 5.4 Pseudo-vecteurs et pseudo-scalaires

A Démontrez, à partir de la transformation de parité (5.6), que l'expression $\bar{\psi}_1 \gamma^5 \psi_2$ est un pseudo-scalaire. ψ_1 et ψ_2 sont deux spineurs de Dirac quelconques ayant la même parité intrinsèque.

B De même, démontrez que la partie spatiale du quadrivecteur $\bar{\psi}_1 \gamma^5 \gamma^\mu \psi_2$ est un vecteur axial, ou pseudo-vecteur.

Problème 5.5 Spineur de Majorana

A Étant donné un spineur de Weyl ψ_L (ou ψ_R), on définit un spineur conjugué de la manière suivante :

$$\psi_L^c = i\sigma_2 \psi_L^* \quad \psi_R^c = i\sigma_2 \psi_R^*$$

Démontrez que, lors des rotations et transformations de Lorentz, ψ_L^c se transforme comme ψ_R et vice-versa : le conjugué d'un spineur gauche se comporte comme un spineur droit et le conjugué d'un spineur droit comme un spineur gauche.

B On définit un *spineur de Majorana* comme formé d'un spineur de Weyl et de son conjugué :

$$\psi_M = \begin{pmatrix} \psi_L \\ \psi_L^c \end{pmatrix} \quad (\text{représentation chirale})$$

Montrez qu'une particule décrite par un spineur de Majorana est nécessairement sa propre anti-particule et que son développement en mode est donné par

$$\psi_M(\mathbf{r}) = \frac{1}{\sqrt{V}} \sum_{\mathbf{p}} \left(c_{\mathbf{p},1} u_{\mathbf{p},1} e^{i\mathbf{p}\cdot\mathbf{r}} + c_{\mathbf{p},2} u_{\mathbf{p},2} e^{i\mathbf{p}\cdot\mathbf{r}} - c_{\mathbf{p},1}^\dagger v_{\mathbf{p},1} e^{-i\mathbf{p}\cdot\mathbf{r}} + c_{\mathbf{p},2}^\dagger v_{\mathbf{p},2} e^{-i\mathbf{p}\cdot\mathbf{r}} \right)$$

Problème 5.6 Constantes de structure

Démontrez que les constantes de structure f_{abc} du groupe $SU(n)$ sont complètement antisymétriques lors de permutations des indices. Indice : utilisez l'antisymétrie du commutateur, la propriété $\text{tr}(T_a T_b) = \frac{1}{2} \delta_{ab}$, ainsi que la propriété cyclique de la trace.

Problème 5.7 État neutre dans le cas de deux couleurs

On peut montrer que le seul état de couleur neutre lorsqu'on combine trois quarks est l'état complètement antisymétrique

$$|\psi\rangle = \frac{1}{\sqrt{6}} (|RGB\rangle + |GBR\rangle + |BRG\rangle - |GRB\rangle - |RBG\rangle - |BGR\rangle) \quad (5.157)$$

Supposons qu'au lieu de trois couleurs, il n'y en ait que deux, et que les baryons soient composés de deux quarks, au lieu de trois, ceci afin de simplifier les choses. Dans ce cas, quelle serait la combinaison complètement antisymétrique ? Démontrez que ce serait aussi la seule combinaison neutre possible. Pour guider votre imagination, notons ces deux couleurs \uparrow et \downarrow , de manière à profiter de l'analogie avec le spin que suggère la symétrie de couleur simplifiée. Cette symétrie est l'invariance par rapport à un mélange des couleurs via la multiplication par une matrice unitaire d'ordre 2.

Problème 5.8 Dérivée covariante du tenseur de Faraday

Comme le tenseur de Faraday $F^{\mu\nu}$ est une matrice, sa dérivée covariante doit être définie légèrement différemment que pour un spineur. Montrer que la définition suivante :

$$\mathcal{D}_\rho F^{\mu\nu} := \partial_\rho F^{\mu\nu} + ig[A_\rho, F^{\mu\nu}]$$

entraîne que $\mathcal{D}_\rho F^{\mu\nu}$ est «covariant de jauge», comme $F^{\mu\nu}$ lui-même, soit

$$\mathcal{D}'_\rho F'^{\mu\nu} = U \mathcal{D}_\rho F^{\mu\nu} U^\dagger$$

Problème 5.9 Champs de couleur

Dans une théorie de jauge comme la chromodynamique (QCD), on peut en principe définir des champs chromo-électriques et chromo-magnétiques, \mathbf{E}^a et \mathbf{B}^a , qui figurent dans le tenseur de Faraday :

$$F^a_{\mu\nu} : \begin{pmatrix} 0 & E_x^a & E_y^a & E_z^a \\ -E_x^a & 0 & -B_z^a & B_y^a \\ -E_y^a & B_z^a & 0 & -B_x^a \\ -E_z^a & -B_y^a & B_x^a & 0 \end{pmatrix} \quad (a = 1, 2, \dots, 8) \quad (5.158)$$

A Écrivez l'expression explicite de ces champs en fonction des potentiels \mathbf{A}^a et Φ^a (la version « couleur » des potentiels vecteur et scalaire).

B Les champs \mathbf{E}^a et \mathbf{B}^a sont-ils invariants de jauge, comme en électrodynamique?

C L'équivalent des équations de Maxwell pour les champs de couleurs s'obtient en remplaçant les dérivées ordinaires par des dérivées covariantes, ce qui se traduit par l'équation suivante :

$$\mathcal{D}_\mu F^{\mu\nu} = J^\nu \quad \text{où on a défini} \quad \mathcal{D}_\mu F^{\mu\nu} := \partial_\mu F^{\mu\nu} + ig[A_\mu, F^{\mu\nu}]$$

Écrivez l'équivalent de la loi de Gauss. En quoi diffère-t-elle de la loi de Gauss pour le champ électromagnétique? Qu'est-ce qui, selon vous, en rend la résolution difficile?

Problème 5.10 Oscillateur 3D et symétrie SU(3)

Considérez le problème d'un oscillateur harmonique en trois dimensions, avec la même constante de ressort dans les trois directions. Les niveaux d'énergie de ce système sont $E_{nmr} = \omega(n + m + r) + E_0$ où ω est la fréquence caractéristique de l'oscillateur, E_0 est l'énergie du fondamental, et les entiers (n, m, r) sont les nombres quantiques entiers associés à chacune des directions de l'espace.

A Quels sont les quatre premiers niveaux d'énergie de ce système, avec leur dégénérescence? Ces dégénérescences peuvent-elles être expliquées uniquement par l'invariance par rotation (c'est-à-dire par des multiplets de moment cinétique orbital l entier), ou existe-t-il une dégénérescence accidentelle?

B Montrez que ce système possède une symétrie $SU(3)$. Cela est facile à voir lorsque le hamiltonien est exprimé en fonction des opérateurs d'échelle.

C Dans cette optique, expliquez la dégénérescence des niveaux en fonction de représentations irréductibles de $SU(3)$.

Problème 5.11 Produits tensoriels de $SU(3)$

A Montrez que le produit tensoriel correspondant à la représentation réductible d'un ensemble de trois quarks est

$$\mathbf{3} \otimes \mathbf{3} \otimes \mathbf{3} = \mathbf{1} \oplus \mathbf{8} \oplus \mathbf{8} \oplus \mathbf{10} \quad (5.159)$$

B Montrez par la même technique qu'un état neutre de couleur à quatre quarks est impossible.

CHAPITRE 6

MODÈLE STANDARD DE L'INTERACTION ÉLECTROFAIBLE

A Interaction faible

6.A.1 Émission bêta

La première manifestation de l'interaction faible en physique subatomique fut l'émission bêta, c'est-à-dire la désintégration d'un noyau instable avec émission d'un électron. L'émission bêta a longtemps été un mystère, du fait que, contrairement à l'émission alpha ou gamma, l'énergie du rayon bêta n'est pas constante, mais suit plutôt un spectre continu (voir Fig. 6.1). Ce comportement est incompréhensible si on suppose que la désintégration n'implique que deux produits : l'électron et le noyau final. En effet, la conservation de l'énergie et de la quantité de mouvement entraîne forcément que, dans une désintégration à deux corps, les deux produits ont des énergies bien définies dans le référentiel du noyau initial. Un paradoxe tout aussi important est la non-conservation du moment cinétique : en effet, les noyaux initial et final ont tous les deux un spin entier ou tous les deux un spin demi-entier, alors que l'électron a un spin $\frac{1}{2}$, ce qui est incompatible avec les règles d'addition du moment cinétique.

La solution à ce paradoxe fut proposée par Pauli en 1930 : la désintégration bêta est en réalité un processus à trois corps, dont l'un est une particule électriquement neutre de spin $\frac{1}{2}$ et de très faible masse que Fermi appela *neutrino*. Cette particule est normalement non détectée, car elle n'interagit que très faiblement avec les nucléons et les électrons. Les vies moyennes relativement longues associées à l'émission bêta (par rapport à l'émission gamma) s'expliquent justement par la faiblesse de l'interaction responsable du processus. Cette interaction, qui n'a rien à voir avec la force nucléaire, est justement connue sous le nom d'*interaction faible*. La répartition de l'énergie entre l'électron et le neutrino dépend de l'angle entre les deux particules, qui est variable. Ceci explique le spectre d'énergie continu de l'électron. Il faudra attendre les années 1950 avant que des signes plus tangibles du neutrino en confirment l'existence.

Prob. 1.9

Sec. 6.E

Une forme simple d'émission bêta est la désintégration du neutron libre :

$$n \rightarrow p + e^- + \bar{\nu}_e \quad (\tau = 886 \text{ s}) \quad (6.1)$$

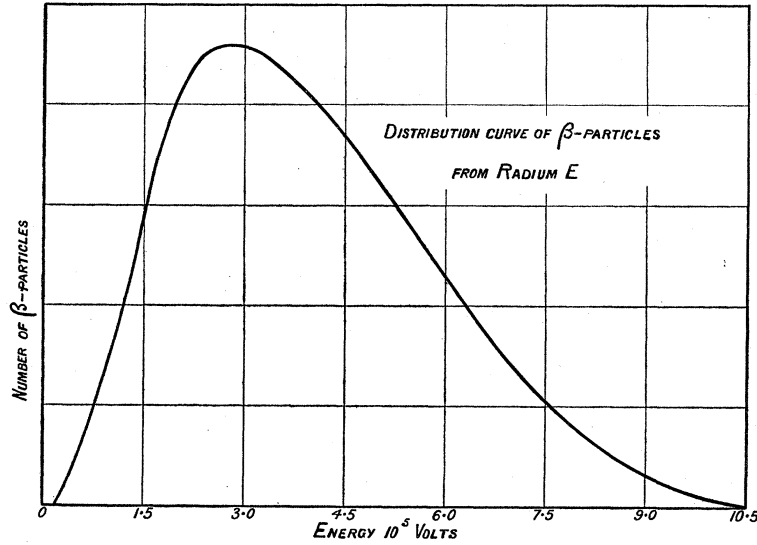


FIGURE 6.1

Spectre d'énergie des rayons bêta (électrons) émis par le radium. Tiré de C.D. Ellis and WA Wooster, Proceedings of the Royal Society of London. Series A **117** (1927) 109123.

Ce processus implique 4 particules : le neutron, le proton, l'électron et le neutrino. Un processus équivalent peut se produire lors d'une réaction, par exemple

$$p + e^- \leftrightarrow n + \nu_e \quad (6.2)$$

Lorsqu'une particule participe à un processus dans un état initial, un autre processus est possible dans lequel l'antiparticule existe dans l'état final, ou vice-versa. Ainsi, la réaction $n + \nu_e \rightarrow p + e^-$ correspond à la désintégration $n \rightarrow p + e^- + \bar{\nu}_e$, où le neutrino initial est devenu un antineutrino dans l'état final. De même, ce pourrait être l'électron qui change de rôle et on pourrait en principe observer la désintégration du proton

$$p \rightarrow n + e^+ + \nu_e \quad (6.3)$$

sauf, évidemment, que la masse du neutron est plus élevée que celle du proton et donc ce processus est impossible par conservation de l'énergie. Par contre, il se produit dans certains noyaux instables, dont la structure altère le bilan énergétique, et est dans ce cas appelé *émission β^+* .

6.A.2 Première théorie de l'interaction faible

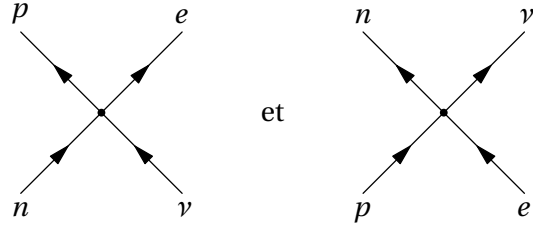
La première théorie de l'interaction faible a été proposée par E. FERMI en 1933. Elle est basée sur le lagrangien d'interaction suivant :

$$\mathcal{L}_I = 2\sqrt{2}G_F \{(\bar{p}\gamma^\mu n)(\bar{e}\gamma_\mu \nu) + (\bar{n}\gamma^\mu p)(\bar{\nu}\gamma_\mu e)\} \quad (6.4)$$

où nous avons défini $\gamma_\mu = g_{\mu\nu}\gamma^\nu$. Dans cette expression, les symboles p , n , e et ν désignent des spineurs de Dirac décrivant respectivement un proton, un neutron, un électron et un neutrino. Le

deuxième terme est le conjugué complexe du premier. Cette théorie fut la première où le concept de création et d'annihilation de particules matérielles fit son apparition.

Ce lagrangien d'interaction mène au vertex suivant :



$$(6.5)$$

De manière équivalente, on pourrait remplacer le proton par un quark u et le neutron par un quark d (les deux autres quarks du nucléon ne jouant aucun rôle dans le processus). La désintégration β s'explique dans cette théorie par un diagramme de Feynman simple avec un neutron comme état initial et un proton, un électron et un antineutrino dans l'état final. La constante de couplage G_F , appelée constante de couplage de Fermi, a les unités d'une masse inverse au carré, et vaut

$$G_F = 1,16639 \times 10^{-5} \text{ GeV}^{-2} \quad (6.6)$$

La théorie de Fermi, cependant, ne brise pas la parité et donc ne peut pas bien représenter tous les processus causés par l'interaction faible, en particulier les processus dans lesquels la polarisation des particules n'est pas moyennée.

6.A.3 Brisure de la parité par l'interaction faible

L'une des propriétés les plus remarquables de l'interaction faible est qu'elle ne respecte pas la symétrie de parité. Cette propriété est apparue premièrement dans la désintégration (faible) des kaons. Les kaons peuvent se désintégrer parfois en des produits ayant une parité paire, parfois en des produits ayant une parité impaire, ce qui contredit la loi de conservation de la parité. T.D. LEE et C.N. YANG ont proposé que l'interaction faible brisait peut-être la parité. Une expérience menée par Mme WU en 1956 confirma cette hypothèse.

Dans cette expérience, on étudie la désintégration d'un noyau de ^{60}Co en présence d'un champ magnétique :



Le champ magnétique permet d'imposer au noyau de cobalt une orientation fixe de son spin ¹.

On pourrait ensuite détecter les électrons émis dans des directions opposées, comme indiqué sur le schéma de la fig. 6.2 : deux détecteurs comptent les électrons et reçoivent des intensités I_1 et I_2 (un seul électron est émis par désintégration, bien sûr, mais on compte ici les électrons émis par un très grand nombre de désintégrations). Si on effectue une rotation de 180 degrés, suivie d'une

1. Pourvu que la température soit suffisamment basse pour que l'orientation du spin ne soit pas désordonnée par les fluctuations thermiques. Ceci nécessite des températures de l'ordre de 10 mK et un processus de désaimantation adiabatique.

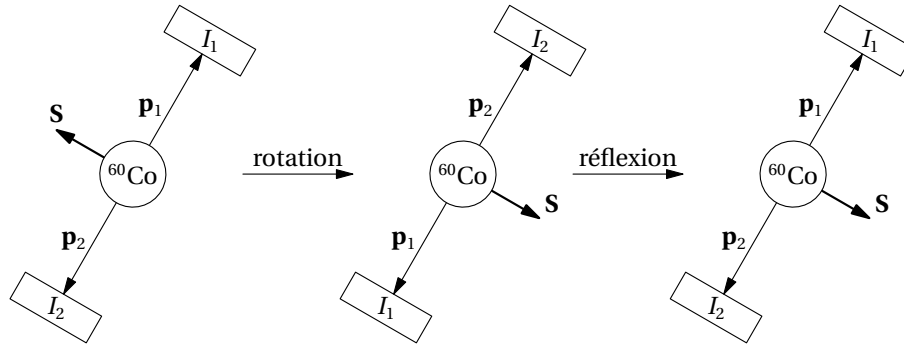
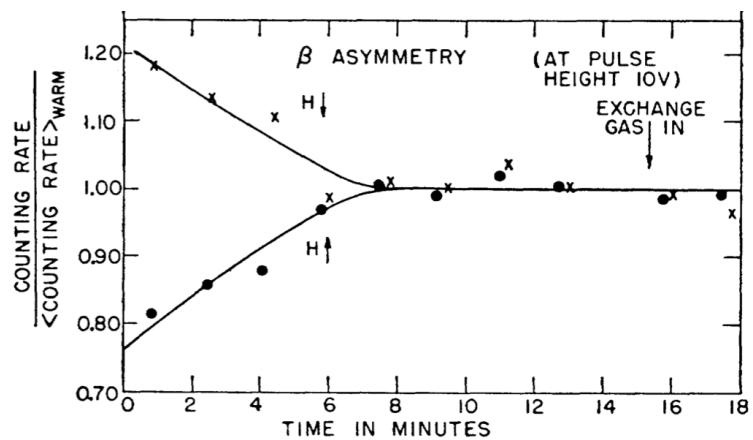


FIGURE 6.2

Schéma de l'expérience de Wu sur la violation de la parité. À gauche, le schéma direct; au centre, celui obtenu par rotation de 180 degrés; à droite, celui obtenu du précédent par inversion de l'espace. On voit que l'inversion de l'espace et la rotation reviennent à inverser le spin.

FIGURE 6.3

Résultats de l'expérience de Wu. Avant que le système nucléaire ne se réchauffe, le nombre de rayons bêta détectés dépend de l'orientation du spin nucléaire. Illustration tirée de : C.S. Wu et al., *Phys. Rev.* **105**, 1413 (1957). (Tous droits réservés)



inversion de l'espace (comme indiqué), alors les deux détecteurs sont inchangés, tandis que le spin est inversé. Si la parité était conservée, alors les deux détecteurs devraient recevoir la même intensité d'électrons ($I_1 = I_2$), ou encore chaque détecteur devrait recevoir le même flux d'électrons, quelle que soit l'orientation du spin des noyaux, car les deux orientations seraient indiscernables. En pratique, on utilise un seul détecteur et on procède à l'inversion du spin en changeant le sens du champ magnétique appliqué. On trouve que l'intensité I_1 n'est pas la même avant et après le changement de sens du champ magnétique (voir Fig.6.3). Donc la parité n'est pas conservée par l'interaction faible.

6.A.4 Théorie V-A

Feynman et Gell-Mann ont formé les quadricourants à caractère polaire et axial suivants :

$$\bar{p}\gamma^\mu n \quad \bar{p}\gamma^\mu\gamma^5 n \quad \bar{e}\gamma^\mu \nu \quad \bar{e}\gamma^\mu\gamma^5 \nu \quad (6.8)$$

et ont proposé le lagrangien suivant pour l'interaction faible :

$$\mathcal{L} = \frac{G_F}{\sqrt{2}}(\bar{p}\gamma^\mu n - \bar{p}\gamma^\mu\gamma^5 n)(\bar{e}\gamma_\mu \nu - \bar{e}\gamma_\mu\gamma^5 \nu) + \text{c.h.} \quad (6.9)$$

appelé communément V-A, car il provient de la soustraction d'un vecteur axial (A) d'un vecteur polaire (V). Ce lagrangien peut aussi être vu comme un couplage identique à celui qui apparaît dans la théorie de Fermi, mais seulement avec la composante gauche du neutron et du neutrino :

$$\mathcal{L} = 2\sqrt{2}G_F(\bar{p}\gamma^\mu n_L)(\bar{e}\gamma_\mu \nu_L) + \text{c.h.} \quad (6.10)$$

Cette théorie de Feynman et Gell-Mann corrige les imperfections de la théorie de Fermi, c'est-à-dire qu'elle contient une brisure explicite de la parité, et est conforme à tous les phénomènes impliquant l'interaction faible connus à l'époque.

Insuffisance de la théorie V-A La théorie V-A, comme la théorie de Fermi originale, ne peut pas être considérée comme une théorie fondamentale de l'interaction faible. La raison en est qu'elle n'est pas renormalisable. Autrement dit, tous les calculs perturbatifs qui sont faits au-delà de l'ordre le plus bas n'ont pas de sens. L'une des conséquences de ce fait, assez simple à démontrer, est la violation de l'unitarité. Expliquons. Considérons la section efficace pour la diffusion par interaction faible entre un électron et un neutrino. Ce processus doit être très peu probable. Dans le référentiel du centre d'impulsion des deux particules, la section efficace doit, au premier ordre de la théorie des perturbations, être proportionnelle à G_F^2 (car G_F apparaît linéairement dans le diagramme de Feynman associé et donc dans l'amplitude de diffusion). Mais l'amplitude de diffusion est sans unité, alors que G_F a les unités d'une masse inverse au carré. Il faut donc que l'amplitude de diffusion soit proportionnelle à une quantité ayant les unités de la masse au carré, soit la masse invariante de la réaction (la seule possibilité). Donc l'amplitude croît comme le carré de la masse invariante, et cette dernière quantité est proportionnelle à l'énergie de l'électron incident. Ceci signifie que la section efficace va croître sans borne avec l'énergie, ce qui est absurde. À un moment donné, la probabilité de collision d'un électron avec un neutrino sera plus grande que

un! C'est ce qu'on appelle une violation de l'unitarité, en somme une non-conservation de la probabilité. Le problème est que les ordres supérieurs de la théorie des perturbations devraient être inclus, et qu'ils ne peuvent l'être parce que la théorie n'est pas renormalisable.

Résumons ici les propriétés les plus notables de l'interaction faible :

1. Elles affectent toutes les particules matérielles connues.
2. Leur portée est très courte, et leur intensité est très faible.
3. Elles peuvent changer la saveur des quarks. Par exemple, les particules comportant un quark s se désintègrent par interaction faible.
4. Elles ne respectent pas l'invariance par inversion de l'espace (parité).

B Théorie de jauge $SU(2) \times U(1)$

La solution proposée aux problèmes de la théorie V-A est de la considérer comme une théorie effective, une approximation d'une théorie plus fondamentale qui serait, elle, renormalisable. Dès 1961, GLASHOW proposa une théorie incomplète qui unifiait l'interaction électromagnétique et l'interaction faible en une seule interaction *électrofaible*. En 1967, WEINBERG et SALAM ont proposé qu'il devait s'agir d'une théorie de jauge non abélienne, mais que l'état fondamental (le vide) ne serait pas invariant de jauge, ce qu'on appelle une brisure spontanée de la symétrie. En 1971, T'HOOFT et VELTMAN démontrèrent que cette théorie compliquée est renormalisable. Glashow, Weinberg et Salam reçurent le prix Nobel en 1979 pour l'élaboration de cette théorie; t'Hooft et Veltman le reçurent une vingtaine d'années plus tard pour avoir démontré que ladite théorie était cohérente. Nous allons maintenant décrire cette théorie.

La théorie électrofaible est une théorie de jauge basée sur le produit $SU(2) \times U(1)$. C'est donc qu'elle comporte trois champs de jauge pour $SU(2)$ (le nombre de générateurs), qu'on notera A_μ^a ($a = 1, 2, 3$) et un champ de jauge pour $U(1)$, qu'on notera B_μ . Ce champ de jauge $U(1)$ se comporte en principe de la même manière que le champ électromagnétique, sauf qu'il ne se couple pas à la charge électrique, mais à une autre quantité Y appelée *hypercharge faible*². Par analogie avec le spin, l'espace abstrait sur lequel agit le groupe $SU(2)$ est appelé espace d'*isospin faible*.

Les fermions fondamentaux qui se couplent à ces champs de jauge doivent occuper des représentations bien précises de $SU(2)$ et avoir une valeur précise de l'hypercharge Y . Dans la théorie électrofaible, tous les fermions sont sans masse au départ (la masse est générée plus tard, par le mécanisme de Higgs décrit plus bas). Donc les composantes droite et gauche des fermions sont indépendantes à ce stade, et peuvent appartenir à des représentations différentes du groupe $SU(2)$, ce qui permet d'introduire une violation maximale de la parité.

2. Ne pas confondre avec l'hypercharge introduite dans la classification des hadrons, qui est une combinaison de l'étrangeté et de la charge électrique.

Commençons par traiter le cas d'une seule famille de particules élémentaires, soit les quarks u et d , l'électron et le neutrino électronique. Les champs associés à ces particules sont agencés comme suit :

$$\ell_L = \begin{pmatrix} \nu_L \\ e_L \end{pmatrix} \quad q_L = \begin{pmatrix} u_L \\ d_L \end{pmatrix} \quad e_R, u_R, d_R \quad (6.11)$$

où on remarque deux doublets d'isospin faible, formés des composantes gauches des leptons et des quarks, respectivement. Les composantes droites forment des singulets d'isospin faible, ce qui signifie qu'ils n'interagissent pas avec le champ de jauge $SU(2)$. Ces différents doublets et singulets ont les valeurs suivantes de l'hypercharge :

champ	Y	$Q = \frac{1}{2}Y + T_3$
$\ell_L = (\nu, e)_L$	-1	$(0, -1)$
$q_L = (u, d)_L$	$\frac{1}{3}$	$(\frac{2}{3}, -\frac{1}{3})$
e_R	-2	-1
u_R	$\frac{4}{3}$	$\frac{2}{3}$
d_R	$-\frac{2}{3}$	$-\frac{1}{3}$

On a aussi indiqué pour chaque champ la valeur de la charge électrique, qui est une combinaison de l'hypercharge et du troisième générateur de $SU(2)$.

Le lagrangien de la théorie est simple : l'action de Dirac s'applique à tous les champs, la dérivée covariante remplaçant la dérivée ordinaire :

$$\mathcal{D}_\mu = \partial_\mu + i g A_\mu^a T^a + i g' \frac{1}{2} Y B_\mu \quad (6.12)$$

où les T^a sont les générateurs de $SU(2)$ dans la représentation considérée, et où le facteur $\frac{1}{2}$ devant B_μ est conventionnel : il affirme simplement que le champ B_μ se couple à l'hypercharge divisée par deux, et non à l'hypercharge directement. Les constantes de couplage g et g' sont distinctes, car elles n'ont pas besoin d'être identiques pour garantir l'invariance de jauge de la théorie. Par exemple, sur le doublet de leptons gauches, on a

$$\mathcal{D}_\mu \ell_L = \left(\partial_\mu + i g \frac{1}{2} A_\mu^a \sigma^a - i \frac{1}{2} g' B_\mu \right) \ell_L \quad (6.13)$$

où les générateurs de $SU(2)$ sont les matrices de Pauli (divisées par 2), alors que pour l'électron droit, on a simplement

$$\mathcal{D}_\mu e_R = (\partial_\mu - i g' B_\mu) e_R \quad (6.14)$$

où les générateurs de $SU(2)$ sont nuls (correspondant à la représentation triviale de $SU(2)$, où tous les éléments du groupe sont associés au nombre 1).

Aux lagrangiens des fermions, on doit ajouter ceux des champs de jauge :

$$\mathcal{L}_{AB} = -\frac{1}{4} F_{\mu\nu}^a F^{a\mu\nu} - \frac{1}{4} G_{\mu\nu} G^{\mu\nu} \quad (6.15)$$

où $F_{\mu\nu}^a$ est le tenseur de Faraday du champ de jauge $SU(2)$ A_μ^a , alors que $G_{\mu\nu}$ est celui du champ de jauge abélien B_μ :

$$F_{\mu\nu}^a = \partial_\mu A_\nu^a - \partial_\nu A_\mu^a - g \varepsilon_{abc} A_\mu^b A_\nu^c \quad G_{\mu\nu} = \partial_\mu B_\nu - \partial_\nu B_\mu \quad (6.16)$$

Notons que les constantes de structure de SU(2) sont simplement $f_{abc} = \varepsilon_{abc}$ d'après les relations de commutation des matrices de Pauli divisées par deux, qui sont celles des composantes du moment cinétique.

Si cette théorie en restait là, alors nous aurions affaire à un ensemble de particules sans masse (oublions la QCD et le confinement pour le moment) interagissant via deux champs de jauge distincts associés à quatre bosons de spin 1 similaires au photon. À la différence de l'électrodynamique quantique, les trois bosons associés aux champs SU(2) interagissent entre eux et forment ensemble une théorie non linéaire. Le lagrangien total décrivant les bosons de jauge et les fermions de la première famille aurait l'expression suivante :

$$\begin{aligned} \mathcal{L} = & -\frac{1}{4}F_{\mu\nu}^a F^{a\mu\nu} - \frac{1}{4}G_{\mu\nu} G^{\mu\nu} \\ & + i\bar{\ell}_L \gamma^\mu \left(\partial_\mu + i g \frac{1}{2} A_\mu^a \sigma^a - i \frac{1}{2} g' B_\mu \right) \ell_L + i\bar{q}_L \gamma^\mu \left(\partial_\mu + i g \frac{1}{2} A_\mu^a \sigma^a + i \frac{1}{6} g' B_\mu \right) q_L \\ & + i\bar{e}_R \gamma^\mu \left(\partial_\mu - i g' B_\mu \right) e_R + i\bar{u}_R \gamma^\mu \left(\partial_\mu + i \frac{2}{3} g' B_\mu \right) u_R + i\bar{d}_R \gamma^\mu \left(\partial_\mu - i \frac{1}{3} g' B_\mu \right) d_R \end{aligned} \quad (6.17)$$

La réalité est plus complexe, car l'interaction faible ont une portée très courte et les trois bosons en question n'ont été découverts qu'en 1983 et ont des masses de l'ordre de 80–90 GeV. C'est la brisure spontanée de l'invariance de jauge qui explique cet état de fait.

C Mécanisme de Higgs

6.C.1 Champ de Higgs

Le mécanisme de Higgs modifie le contenu physique de la théorie électrofaible décrite ci-haut en conférant une masse à trois des quatre bosons de jauge et en donnant une masse non nulle aux fermions. Il se base sur la présence dans la théorie d'un champ scalaire Φ , en plus des fermions décrits plus haut et des champs de jauge. Ce champ scalaire représente un boson de spin zéro, qui forme cependant un doublet complexe d'isospin faible et qui possède une hypercharge $Y = 1$. On peut exprimer ce doublet comme

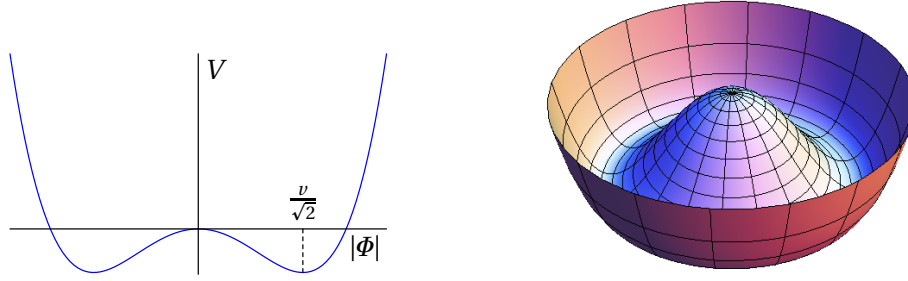
$$\Phi = \begin{pmatrix} \phi^+ \\ \phi^0 \end{pmatrix} \quad (6.18)$$

où l'assignation des charges électriques des membres du doublet correspond bien à $Q = \frac{1}{2}Y + T_3$.

On pourrait, raffinement supplémentaire, ajouter un terme en $|\phi|^4$ à ce lagrangien :

$$\mathcal{L} = (\mathcal{D}_\mu \phi)^* \mathcal{D}^\mu \phi - m^2 |\phi|^2 - \lambda |\phi|^4 \quad (6.19)$$

Dans ce cas, la théorie n'est plus soluble exactement. Le terme en $|\phi|^4$ correspond à une interaction du champ avec lui-même, un peu comme dans la théorie de Fermi de l'interaction faible, sauf que les particules impliquées sont des bosons. Notons que la constante de couplage λ n'a pas d'unité,

**FIGURE 6.4**

Le potentiel de Higgs $V(|\Phi|)$. À gauche, coupe en fonction de $|\Phi|$. À droite, coupe en fonction de $|\Phi|$ et d'une des variables de phase, affectant la forme d'un « chapeau mexicain ».

ce qui entraîne que la théorie est renormalisable et qu'elle constitue donc une théorie physique acceptable en principe.

Pour qu'un champ scalaire puisse aussi interagir avec les champs de jauge A_μ^a de SU(2), il ne doit pas former un singulet de SU(2), mais appartenir à une représentation non triviale de SU(2). La possibilité la plus simple est le doublet décrit plus haut.

Le lagrangien de ce doublet, dit doublet de Higgs, est

$$\mathcal{L}_H = (\mathcal{D}_\mu \Phi)^\dagger (\mathcal{D}^\mu \Phi) - V(\Phi) \quad V(\Phi) = -\mu^2 \Phi^\dagger \Phi + \lambda (\Phi^\dagger \Phi)^2 \quad (6.20)$$

Le premier terme du lagrangien est simplement le lagrangien d'un champ scalaire ordinaire, où les dérivées ordinaires ont été remplacées par des dérivées covariantes pour assurer l'invariance de jauge. Comme Φ est un doublet d'hypercharge 1, on a

$$\mathcal{D}_\mu \Phi = \partial_\mu \Phi + \frac{1}{2} i g A_\mu^a \sigma^a \Phi + \frac{1}{2} i g' B_\mu \Phi \quad (6.21)$$

Le deuxième terme du lagrangien est un potentiel $V(\Phi)$ qui ne dépend que de la grandeur $\Phi^\dagger \Phi = |\phi^+|^2 + |\phi^0|^2$. Comme V ne dépend que de $\Phi^\dagger \Phi$, il est, comme le reste du lagrangien, invariant par transformation de jauge

$$\Phi \rightarrow \Phi' = U e^{i\theta} \Phi \quad (6.22)$$

où U est une matrice appartenant à SU(2) et θ est un facteur de phase associé à la symétrie de jauge U(1) de l'hypercharge. Il faut voir le potentiel V dans un espace à quatre dimensions (Φ possède l'équivalent de quatre composantes réelles) dont seule une coupe dans une direction particulière est illustrée à la fig. 6.4.

6.C.2 Brisure spontanée de la symétrie

Le potentiel $V(|\Phi|)$ a ceci de particulier que, en raison du signe négatif du terme quadratique, son minimum n'est pas situé à $|\Phi| = 0$, mais à une valeur non nulle de $|\Phi|$ qu'on note $v/\sqrt{2}$ (voir fig. 6.4). Ceci

Densité lagrangienne du modèle standard, première famille de fermions élémentaires

$$\begin{aligned}
 \mathcal{L} = & -\frac{1}{4}(\partial_\mu A_\nu^a - \partial_\nu A_\mu^a - g\epsilon_{abc}A_\mu^b A_\nu^c)(\partial^\mu A^{a\nu} - \partial^\nu A^{a\mu} - g\epsilon_{abc}A^{b\mu} A^{c\nu}) \\
 & -\frac{1}{4}(\partial_\mu B_\nu - \partial_\nu B_\mu)(\partial^\mu B^\nu - \partial^\nu B^\mu) \\
 & + i\bar{\ell}_L \gamma^\mu \left(\partial_\mu + i g \frac{1}{2} A_\mu^a \sigma^a - i \frac{1}{2} g' B_\mu \right) \ell_L + i\bar{q}_L \gamma^\mu \left(\partial_\mu + i g \frac{1}{2} A_\mu^a \sigma^a + i \frac{1}{6} g' B_\mu \right) q_L \\
 & + i\bar{e}_R \gamma^\mu (\partial_\mu - i g' B_\mu) e_R + i\bar{u}_R \gamma^\mu (\partial_\mu + i \frac{2}{3} g' B_\mu) u_R + i\bar{d}_R \gamma^\mu (\partial_\mu - i \frac{1}{3} g' B_\mu) d_R \\
 & + (\partial_\mu \Phi^\dagger - \frac{1}{2} i g A_\mu^a \sigma^a \Phi^\dagger - \frac{1}{2} i g' B_\mu \Phi^\dagger)(\partial^\mu \Phi + \frac{1}{2} i g A^{a\mu} \sigma^a \Phi + \frac{1}{2} i g' B^\mu \Phi) + \mu^2 \Phi^\dagger \Phi - \lambda(\Phi^\dagger \Phi)^2
 \end{aligned}
 \tag{6.23}$$

signifie que la valeur du champ Φ n'est pas nulle dans l'état fondamental, mais telle que $|\Phi| = v/\sqrt{2}$.³ Dans l'état fondamental, le système choisit une direction dans cet espace quadridimensionnel défini par les quatre composantes réelles de Φ . Ce choix se fait de manière accidentelle, par les conditions initiales de l'Univers. C'est un caractère contingent de la théorie sur lequel il est difficile de spéculer.

Le phénomène par lequel une symétrie est violée par l'état fondamental du système et non pas le lagrangien (ou le hamiltonien) porte le nom de *brisure spontanée de la symétrie*. Ce mécanisme intervient couramment dans les théories d'unification des particules et, surtout, en physique de la matière condensée. Une bonne partie de la physique des changements de phase repose sur la notion de brisure spontanée de symétrie. Il est important de comprendre, cependant, qu'une brisure spontanée de la symétrie ne peut se produire que dans des systèmes comportant un très grand nombre de degrés de liberté (c'est-à-dire dans la limite thermodynamique).

Considérons à cet effet un potentiel en forme de « chapeau mexicain » (voir fig. 6.4). Si on place une particule (un boson) dans ce puits de potentiel, l'état fondamental sera tout à fait symétrique et la densité de probabilité associée aura la forme d'un anneau, avec un maximum près du minimum du potentiel V . En fait, si on dépose une particule à un endroit précis du potentiel, cet état préparé ne serait pas un état propre du hamiltonien, et évoluerait de manière complexe, comme un paquet d'ondes, en s'élargissant et se compactant à tour de rôle dans la direction azimutale. Une faible interaction avec l'environnement ferait que l'état fondamental finirait par dominer et la fonction d'onde deviendrait de plus en plus symétrique : la symétrie serait restaurée dans un temps fini. Supposons maintenant qu'on dépose un grand nombre de bosons indiscernables dans ce puits de potentiel et qu'en plus une légère interaction attractive les pousse à s'agglomérer. Les bosons formeraient alors un « amas » à un endroit donné, par condensation de Bose-Einstein et la position de cet amas violerait la symétrie du potentiel. On peut montrer alors que le temps requis pour que cet amas s'étale sur le pourtour du potentiel et se transforme en un état complètement symétrique tend vers l'infini : la symétrie est effectivement brisée. C'est là l'essence du phénomène de brisure spontanée de la symétrie. On pourrait aussi le décrire à l'aide d'un système de spins en interaction, dans le cadre d'un changement de phase magnétique. Le phénomène est universel.

3. Du moins, dans l'approximation classique. Des corrections quantiques modifieraient légèrement cette valeur.

Le potentiel de Higgs introduit de manière *ad hoc* la brisure spontanée de la symétrie de jauge $SU(2) \times U(1)$ dans le modèle standard. Comme une théorie des champs comporte un nombre infini de degrés de liberté, la chose est permise. Nous allons supposer, en première approximation, que le champ de Higgs Φ prend une valeur moyenne non nulle et montrer comment les champs de jauge associés à la symétrie brisée deviennent massifs. Ils sont en quelque sorte « alourdis » par leur interaction avec un champ qui possède une valeur constante, comme nous allons maintenant le voir en détail.

Supposons donc que le champ de Higgs est exactement égal à sa valeur dans l'état fondamental, c'est-à-dire

$$\Phi = \begin{pmatrix} 0 \\ v/\sqrt{2} \end{pmatrix} \quad (6.24)$$

et négligeons les fluctuations du champ Φ autour de cette valeur constante. Nous avons manifestement choisi une direction dans l'espace interne du doublet, en supposant que la valeur de ϕ^+ dans le vide est nulle – ce qui signifie a posteriori que le vide n'est pas chargé électriquement et donc que l'invariance de jauge électromagnétique est préservée – et que la valeur de ϕ^0 est réelle.

Génération des masses des bosons intermédiaires Voyons maintenant ce que produit le lagrangien du doublet de Higgs quand on substitue cette valeur de Φ . Toutes les dérivées sont nulles, bien sûr, mais il reste les couplages aux champs de jauge, qui deviennent

$$\mathcal{D}_\mu \Phi \rightarrow \frac{v}{\sqrt{2}} \left(\frac{1}{2} i g A_\mu^a \sigma^a + \frac{1}{2} i g' B_\mu \right) \begin{pmatrix} 0 \\ 1 \end{pmatrix} = \frac{i v}{2\sqrt{2}} \begin{pmatrix} g A_\mu^1 - i g A_\mu^2 \\ g' B_\mu - g A_\mu^3 \end{pmatrix} \quad (6.25)$$

Le lagrangien du doublet devient alors

$$\begin{aligned} \mathcal{L}_H &\rightarrow \frac{v^2}{8} \begin{pmatrix} g A_\mu^1 - i g A_\mu^2 \\ g' B_\mu - g A_\mu^3 \end{pmatrix}^\dagger \begin{pmatrix} g A^{1\mu} - i g A^{2\mu} \\ g' B^\mu - g A^{3\mu} \end{pmatrix} \\ &= \frac{v^2}{8} \left\{ g^2 (A_\mu^1 A^{1\mu} + A_\mu^2 A^{2\mu}) + (g' B_\mu - g A_\mu^3)(g' B^\mu - g A^{3\mu}) \right\} \end{aligned} \quad (6.26)$$

On voit que des termes quadratiques dans les champs de jauge sont générés, ce qui correspond à conférer à ces champs une masse. En effet, un terme quadratique (de signe négatif) dans le lagrangien du champ scalaire est associé à une masse de ce champ. Ceci vaut aussi pour chacune des composantes d'un champ vectoriel, comme pour un champ scalaire. Mais ce ne sont pas tous les champs de jauge qui acquièrent une masse dans ce mécanisme. En fait, dans toute théorie des champs, l'ensemble des termes quadratiques dans les champs forment une forme quadratique qu'il est utile de considérer dans la base de ses états propres; dit autrement : il vaut mieux utiliser les combinaisons des champs qui rendent cette forme quadratique diagonale.

Les termes du lagrangien (6.23) qui sont effectivement quadratiques en A et B sont alors les suivants :

$$\begin{aligned} \mathcal{L}_{\text{quad.}} &= -\frac{1}{4} (\partial_\mu A_\nu^a - \partial_\nu A_\mu^a) (\partial^\mu A^{a\nu} - \partial^\nu A^{a\mu}) - \frac{1}{4} (\partial_\mu B_\nu - \partial_\nu B_\mu) (\partial^\mu B^\nu - \partial^\nu B^\mu) \\ &\quad + \frac{v^2}{8} \left\{ g^2 (A_\mu^1 A^{1\mu} + A_\mu^2 A^{2\mu}) + (g' B_\mu - g A_\mu^3)(g' B^\mu - g A^{3\mu}) \right\} \end{aligned} \quad (6.27)$$

Dans le but de préserver la normalisation des tenseurs de Faraday qui figurent dans cette expression, on utilise une combinaison des champs de jauge qui constitue en fait une rotation dans l'espace des champs de jauge :

$$\begin{aligned} Z_\mu &= \cos \theta_W A_\mu^3 - \sin \theta_W B_\mu \\ A_\mu &= \sin \theta_W A_\mu^3 + \cos \theta_W B_\mu \end{aligned} \quad (6.28)$$

où l'angle de Weinberg θ_W est défini par

$$\sin \theta_W = \frac{g'}{\sqrt{g^2 + g'^2}} \quad \cos \theta_W = \frac{g}{\sqrt{g^2 + g'^2}} \quad \tan \theta_W = \frac{g'}{g} \quad (6.29)$$

On utilise en outre des combinaisons complexes des champs A^1 et A^2 :

$$W_\mu = \frac{1}{\sqrt{2}} (A_\mu^1 - i A_\mu^2) \quad (6.30)$$

En fonction de ces combinaisons, les termes quadratiques du lagrangien deviennent

$$\begin{aligned} \mathcal{L}_{\text{quad.}} &= -\frac{1}{4} (\partial_\mu A_\nu - \partial_\nu A_\mu) (\partial^\mu A^\nu - \partial^\nu A^\mu) - \frac{1}{4} (\partial_\mu Z_\nu - \partial_\nu Z_\mu) (\partial^\mu Z^\nu - \partial^\nu Z^\mu) \\ &\quad - \frac{1}{2} (\partial_\mu W_\nu - \partial_\nu W_\mu)^* (\partial^\mu W^\nu - \partial^\nu W^\mu) \\ &\quad + m_W^2 W^{\mu*} W_\mu + \frac{1}{2} m_Z^2 Z^\mu Z_\mu \end{aligned} \quad (6.31)$$

où on a défini les masses

$$m_W = \frac{vg}{2} \quad m_Z = \frac{v}{2} \sqrt{g^2 + g'^2} \quad \frac{m_W}{m_Z} = \frac{g}{\sqrt{g^2 + g'^2}} = \cos \theta_W \quad (6.32)$$

Ce lagrangien décrit des particules vectorielles (c'est-à-dire de spin 1) :

1. Le photon (symbole γ), décrit par le champ électromagnétique A_μ , sans masse.
2. Le W^\pm , particule chargée, de masse m_W , décrite par le champ complexe W_μ .
3. Le Z , particule neutre, de masse m_Z , décrite par le champ réel Z_μ .

Les valeurs observées des masses sont :

$$m_W = 80,398 \text{ GeV} \quad m_Z = 91,19 \text{ GeV} \quad \sin^2 \theta_W \approx 0,2311 \quad (6.33)$$

La combinaison A_μ , qui est en fait le champ électromagnétique, demeure sans masse.

Pourquoi peut-on interpréter m_W et m_Z comme étant les masses des quanta des champs W_μ et Z_μ ? La réponse se trouve dans la solution du problème 2.1. Remarquons que le terme de masse ci-dessus a la même forme que celui qui apparaît dans le lagrangien d'un champ scalaire (ou d'un champ scalaire complexe, sans le préfacteur $\frac{1}{2}$), sauf que ce terme a ici le signe opposé, ce qui est adéquat pour un champ vectoriel : trois des quatre termes de la contraction $Z_\mu Z^\mu$ ont un signe négatif.

L'hypothèse que la masse d'un boson de jauge puisse apparaître suite à la brisure spontanée de la symétrie de jauge a été formulée indépendamment par plusieurs auteurs : Brout et Englert⁴, Higgs⁵, Guralnik, Hagen et Kibble⁶. Pour des raisons difficiles à cerner, c'est le nom de Peter Higgs qui reste associé à ce mécanisme et au boson associé. Signalons aussi que la même idée générale a été proposée par P.W. Anderson un peu auparavant, dans le contexte de la matière condensée.⁷

Modes de Goldstone et polarisations des bosons massifs

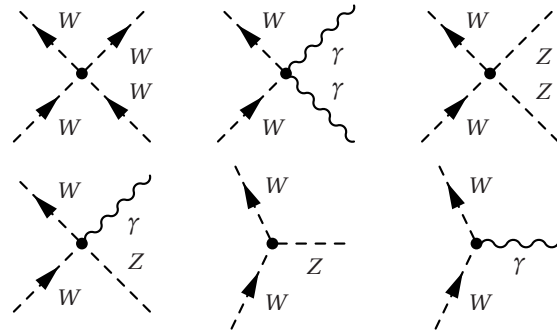
En général, un système physique qui subit la brisure spontanée d'une symétrie continue présente ce qu'on appelle des *modes de Goldstone* : ce sont des excitations dont l'énergie s'annule avec le vecteur d'onde (par exemple $E \propto |\mathbf{k}|$) et qui correspondent à une variation lente du champ qui brise la symétrie. Dans le cas qui nous occupe, une configuration du champ Φ qui varie lentement aurait une énergie d'autant plus basse que la variation est lente.

Cependant, si la symétrie brisée est une symétrie de jauge, ces modes de Goldstone se manifestent d'une manière particulière : ils sont en quelque sorte incorporés au champ de jauge devenu massif et fournissent les degrés de liberté nécessaires à l'apparition d'une troisième polarisation de spin des bosons de jauge. En effet, un boson de jauge sans masse, comme le photon, ne possède que deux polarisations (deux hélicités), malgré qu'il représente une particule de spin 1. Si cette particule, suite à une brisure spontanée de la symétrie de jauge, devient massive, alors l'hélicité n'est plus un bon nombre quantique, et les trois projections de spin associées à $j = 1$ doivent apparaître, car il existe toujours un référentiel dans lequel la particule massive est au repos. L'apparition d'une troisième projection de spin nécessite une augmentation du nombre de degrés de liberté du système, qui provient justement de ces fameux modes de Goldstone.

La seule fluctuation possible du champ Φ autour de sa valeur moyenne est donc dans la direction radiale de Φ , et représente ce qu'on appelle la particule de Higgs (voir le problème 6.2).

FIGURE 6.5

Vertex décrivant les interactions entre champs de jauge dans le modèle standard. Les lignes brisées non orientées représentent le Z ; les lignes brisées orientées le W ; les lignes ondulées le photon. Les vertex à quatre pattes comptent pour deux ordres de la théorie des perturbations (g^2) et les vertex à trois pattes pour un ordre (g).



Interactions entre les bosons de jauge

En fonction des nouveaux champs W , Z et A , les termes non linéaires impliquant les champs de jauge A^a deviennent

$$\mathcal{L}_{\text{N.L.}} = \frac{g^2}{4} (W_\mu^* W_\nu - W_\nu^* W_\mu) (W^{*\mu} W^\nu - W^{*\nu} W^\mu)$$

4. F. Englert et R. Brout, Phys. Rev. Lett., **13**, 321323 (1964).

5. P. W. Higgs, Phys. Rev. Lett., **13**, 508 (1964).

6. G. S. Guralnik, C. R. Hagen, et T. W. B. Kibble, Phys. Rev. Lett. **13**, 585 (1964).

7. P. W. Anderson, Phys. Rev. **130**, 439 (1963).

$$\begin{aligned}
 & -\frac{g^2}{2} \sin^2 \theta_W \left[2A_\mu A^\mu W_\nu^* W^\nu - A^\mu A^\nu (W_\mu^* W_\nu + W_\nu^* W_\mu) \right] \\
 & -\frac{g^2}{2} \cos^2 \theta_W \left[2Z_\mu Z^\mu W_\nu^* W^\nu - Z^\mu Z^\nu (W_\mu^* W_\nu + W_\nu^* W_\mu) \right] \\
 & -\frac{g^2}{2} \sin \theta_W \cos \theta_W \left[4A_\mu Z^\mu W_\nu^* W^\nu - (A^\mu Z^\nu + A^\nu Z^\mu) (W_\mu^* W_\nu + W_\nu^* W_\mu) \right] \\
 & + g \sin \theta_W \operatorname{Im} \left[(\partial_\mu W_\nu - \partial_\nu W_\mu)^* (A^\mu W^\nu - A^\nu W^\mu) + (\partial_\mu A_\nu - \partial_\nu A_\mu) (W^{*\mu} W^\nu - W^{*\nu} W^\mu) \right] \\
 & + g \cos \theta_W \operatorname{Im} \left[(\partial_\mu W_\nu - \partial_\nu W_\mu)^* (Z^\mu W^\nu - Z^\nu W^\mu) + (\partial_\mu Z_\nu - \partial_\nu Z_\mu) (W^{*\mu} W^\nu - W^{*\nu} W^\mu) \right] \quad (6.34)
 \end{aligned}$$

Les termes en g^2 sont quartiques et les termes en g sont cubiques. Les différents bosons de jauge, incluant le photon, interagissent donc entre eux. En particulier, il existe un terme cubique représenté dans les règles de Feynman par un vertex où une particule W émet ou absorbe un photon, ou un Z . Il y a en outre des termes quartiques représentés par des vertex où une particule W émet deux photons, deux Z , ou un photon et un Z . Il y a enfin des termes quartiques n'impliquant que des W . Cependant, le Z n'émet ou n'absorbe pas de photon et vice-versa : ceci signifie que le Z est neutre électriquement, contrairement au W qui se comporte comme une particule chargée, de charge $+1$, avec son antiparticule de charge -1 , représentée par le champ conjugué W^* (nous verrons plus bas comment la charge du W est déterminée). Voir la figure 6.5 pour la représentation diagrammatique de ces termes.

6.C.3 Courants chargé, neutre et électromagnétique

En fonction des nouveaux champs de jauge W , Z et A , le lagrangien d'interaction entre les fermions et les bosons de jauge a une apparence différente. Pour mieux l'exprimer, nous utilisons la notion de *courant* et écrivons les termes d'interaction comme

$$\mathcal{L}_I = -g J^a \cdot A^a - g' J^Y \cdot B \quad (6.35)$$

où nous avons supprimé les indices d'espace-temps pour simplifier les expressions en introduisant la notation $A \cdot B := A_\mu B^\mu$, et où nous avons défini les courants suivants :

$$\begin{aligned}
 J^{a\mu} &= \frac{1}{2} \bar{\ell}_L \gamma^\mu \sigma^a \ell_L + \frac{1}{2} \bar{q}_L \gamma^\mu \sigma^a q_L \\
 J^{Y\mu} &= \frac{1}{2} \sum_\alpha Y_\alpha \bar{f}_\alpha \gamma^\mu f_\alpha
 \end{aligned} \quad (6.36)$$

La somme dans la deuxième équation est prise sur toutes les espèces de fermions f_α présentes, droites et gauches, et Y_α est l'hypercharge de chaque espèce.

Il est plus pratique d'exprimer le lagrangien d'interaction (6.35) en fonction des nouveaux champs de jauge W_μ , Z_μ et A_μ , de la manière suivante :

$$\mathcal{L}_I = -\frac{g}{\sqrt{2}} (J \cdot W + J^* \cdot W^*) - e J_{\text{em}} \cdot A - \frac{g}{\cos \theta_W} Z \cdot J_0 \quad (6.37)$$

L'expression ci-dessus constitue en fait une définition du courant chargé J , du courant neutre J_0 et du courant électromagnétique J_{em} . Ces courants sont des expressions bilinéaires en fermions qui permettent d'identifier les particules qui interagissant directement avec les champs W_μ , Z_μ et A_μ .

Courant chargé On voit facilement que les courants J^1 et J^2 peuvent être réorganisés en composantes complexes $J = J^1 + iJ^2$ et $J^* = J^1 - iJ^2$, ce qui mène à un premier terme d'interaction

$$\begin{aligned}\mathcal{L}_{\text{CC}} &= -\frac{1}{2}g(J + J^*) \cdot A^1 + i\frac{1}{2}g(J - J^*) \cdot A^2 \\ &= -\frac{1}{2}gJ \cdot (A^1 - iA^2) - \frac{1}{2}gJ^* \cdot (A^1 + iA^2) \\ &= -\frac{g}{\sqrt{2}}(J \cdot W + J^* \cdot W^*)\end{aligned}\quad (6.38)$$

Le courant J est appelé *courant chargé* parce qu'il est couplé à un boson de jauge chargé (le W) et parce qu'ils ne sont pas invariants lors d'une transformation de jauge électromagnétique. Développons plus en détail l'expression du courant chargé. On voit immédiatement que

$$J^\mu = \bar{\ell}_L \gamma^\mu \sigma^+ \ell_L + \bar{q}_L \gamma^\mu \sigma^+ q_L \quad \sigma^+ = \frac{1}{2}(\sigma^1 + i\sigma^2) = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} \quad (6.39)$$

ou encore, en explicitant les doublets d'isospin faible,

$$J^\mu = \bar{\nu}_L \gamma^\mu e_L + \bar{u}_L \gamma^\mu d_L \quad (6.40)$$

alors que

$$J^{\mu*} = \bar{e}_L \gamma^\mu \nu_L + \bar{d}_L \gamma^\mu u_L \quad (6.41)$$

Le lagrangien d'interaction impliquant le courant chargé est donc

$$\mathcal{L}_{\text{CC}} = -\frac{g}{\sqrt{2}}(\bar{\nu}_L \gamma^\mu e_L + \bar{u}_L \gamma^\mu d_L)W_\mu + \text{c.h.} \quad (6.42)$$

Comme la première parenthèse décrit un objet de charge -1 , le champ W doit décrire un objet de charge $+1$, afin que le lagrangien soit neutre. Le champ conjugué W^* décrit alors un objet de charge -1 .

Courants neutre et électromagnétique Considérons ensuite la partie restante de l'équation (6.35) :

$$\mathcal{L}_{\text{CN}} = -gJ^3 \cdot A^3 - g'J^Y \cdot B \quad (6.43)$$

Substituons dans cette expression les champs Z et A , en utilisant les relations inverses

$$A^3 = \cos \theta_W Z + \sin \theta_W A \quad B = -\sin \theta_W Z + \cos \theta_W A \quad (6.44)$$

On trouve alors

$$\begin{aligned}\mathcal{L}_{\text{CN}} &= -gJ^3 \cdot (\cos \theta_W Z + \sin \theta_W A) - g'J^Y \cdot (-\sin \theta_W Z + \cos \theta_W A) \\ &= -A \cdot (g \sin \theta_W J^3 + g' \cos \theta_W J^Y) - Z \cdot (g \cos \theta_W J^3 - g' \sin \theta_W J^Y)\end{aligned}\quad (6.45)$$

Notons cependant que $g \sin \theta_W = g' \cos \theta_W$, de sorte qu'on peut également écrire

$$\mathcal{L}_{\text{CN}} = -g \sin \theta_W A \cdot (J^3 + J^Y) - \frac{g}{\cos \theta_W} Z \cdot (\cos^2 \theta_W J^3 - \sin^2 \theta_W J^Y) \quad (6.46)$$

On définit ensuite la nouvelle constante de couplage et le courant électromagnétiques :

$$e = g \sin \theta_W \quad J_{\text{em}} = J^3 + J^Y \quad (6.47)$$

De plus, comme $\cos^2 \theta_W J^3 = J^3 - J^3 \sin^2 \theta_W$, on écrit simplement

$$\begin{aligned} \mathcal{L}_{\text{CN}} &= -e A \cdot J_{\text{em}} - \frac{g}{\cos \theta_W} Z \cdot (J^3 - \sin^2 \theta_W (J^3 + J^Y)) \\ &= -e A \cdot J_{\text{em}} - \frac{g}{\cos \theta_W} Z \cdot J_0 \end{aligned} \quad (6.48)$$

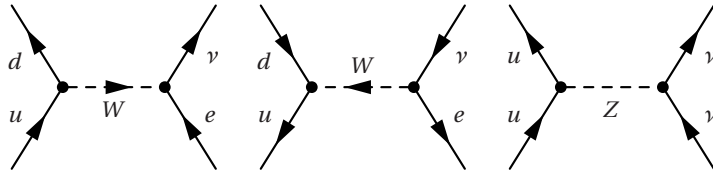
où nous avons défini le courant neutre

$$J_0 = J^3 - \sin^2 \theta_W J_{\text{em}} \quad (6.49)$$

Le quotient $g/\cos \theta_W$ représente la constante de couplage faible avec laquelle le courant neutre interagit avec le champ Z , alors que les courants chargés interagissent avec W avec une constante de couplage g .

FIGURE 6.6

Diagrammes typiques résultants des interactions des bosons W^\pm et Z avec les courants chargés et neutres, respectivement. La direction de la flèche décorant le W est déterminée par la conservation de la charge électrique.



Développons plus en détail le courant J^3 :

$$J^{3\mu} = \frac{1}{2} (\bar{\nu}_L \gamma^\mu \nu_L - \bar{e}_L \gamma^\mu e_L + \bar{u}_L \gamma^\mu u_L - \bar{d}_L \gamma^\mu d_L) \quad (6.50)$$

Alors que le courant J^Y est

$$J^{Y\mu} = \frac{1}{2} \left(-\bar{\nu}_L \gamma^\mu \nu_L - \bar{e}_L \gamma^\mu e_L + \frac{1}{3} \bar{u}_L \gamma^\mu u_L + \frac{1}{3} \bar{d}_L \gamma^\mu d_L \right) - \bar{e}_R \gamma^\mu e_R + \frac{2}{3} \bar{u}_R \gamma^\mu u_R - \frac{1}{3} \bar{d}_R \gamma^\mu d_R \quad (6.51)$$

Le courant électromagnétique J_{em} est alors

$$J_{\text{em}} = J^{3\mu} + J^{Y\mu} = (-\bar{e} \gamma^\mu e + \frac{2}{3} \bar{u} \gamma^\mu u - \frac{1}{3} \bar{d} \gamma^\mu d) \quad (6.52)$$

Dans cette dernière expression, nous avons combiné les parties gauches et droites. Remarquons à cet effet que, pour un spineur de Dirac,

$$\bar{\psi}_L \gamma^\mu \psi_L + \bar{\psi}_R \gamma^\mu \psi_R = \bar{\psi} \gamma^\mu \psi \quad (6.53)$$

comme on le voit facilement en substituant les définitions de ψ_L et ψ_R en fonction de γ^5 et en utilisant l'anticommutation de γ^5 avec les matrices de Dirac.

Des diagrammes de Feynman typiques résultant des lagrangiens \mathcal{L}_{CC} et \mathcal{L}_{CN} représentant l'interaction faible sont illustrés à la figure 6.6. Notons que les courants neutres impliquent des fermions de la même espèce, car ils sont construits à partir de générateurs diagonaux. Par contre, les courants chargés impliquent un changement d'espèce de fermions, à l'intérieur d'un même doublet d'isospin faible.

6.C.4 Relation avec la théorie V-A

Pourquoi l'interaction faible est-elle faible ? Parce que les bosons de jauge associés sont massifs. Le facteur associé au boson virtuel (ou propagateur) est

$$\frac{1}{q^2 - m^2} \quad (6.54)$$

Dans le cas des bosons W et Z , et de processus impliquant des transferts d'énergie et d'impulsion faibles devant m , ce facteur se réduit à $1/m^2$. L'amplitude de diffusion obtenue du diagramme de Feynman est donc de l'ordre g^2/m_W^2 pour les courants chargés, et en fait identique pour le courant neutre, étant donné que $m_Z \cos \theta_W = m_W$.

On pourrait en fait remplacer l'interaction de jauge par une interaction efficace courant-courant qui aurait la forme suivante :

$$\mathcal{L}_{CC}^{\text{eff.}} = -\frac{g^2}{2m_W^2} J_\mu^+ J^{-\mu} \quad \mathcal{L}_{CN}^{\text{eff.}} = -\frac{g^2}{2 \cos^2 \theta_W m_Z^2} J_\mu^0 J_0^\mu = -\frac{g^2}{2m_W^2} J_\mu^0 J_0^\mu \quad (6.55)$$

On peut montrer que cette interaction (du moins la partie courants chargés) coïncide avec l'interaction V-A. La relation entre les constantes de couplage est

$$G_F = \frac{g^2}{4\sqrt{2}m_W^2} = 1,091 \times 10^{-5} \text{ GeV}^{-2} \quad (6.56)$$

La dimension (les unités) de la constante G_F et sa petite valeur trouvent ici une explication naturelle. Notons que si cette valeur ne coïncide pas exactement avec la valeur mesurée de la constante de Fermi ($1,16639 \times 10^{-5} \text{ GeV}^{-2}$), c'est que le résultat ci-dessus est approximatif (théorie des perturbations aux 2^{ème} ordre).

En fait, l'interaction faible n'est faible qu'aux petites énergies. Dans les processus à très haute énergie, impliquant des transferts d'énergie et de quantité de mouvement de l'ordre de la masse m_W ou plus grands, l'interaction faible est d'une force comparable à l'interaction électromagnétique. En fait, à ces grandes énergies, les masses des bosons intermédiaires perdent de leur importance. La constante de couplage $g^2/4\pi = e^2/(4\pi \sin^2 \theta_W) \approx 1/30$ est en fait plus grande que la constante de structure fine électromagnétique, de sorte que ces interactions sont en fait plus intenses que l'interaction électromagnétique, mais restent tout de même dans le régime où la théorie des perturbations est applicable.

D Génération des masses des fermions

6.D.1 Couplage de Yukawa

Jusqu'ici, dans notre traitement de la théorie électrofaible, les fermions sont toujours sans masse. Il est possible de leur conférer une masse par le mécanisme de Higgs, comme pour les bosons de jauge. Il suffit pour cela de postuler que le doublet de Higgs Φ est couplé aux fermions par un terme d'interaction simple, dit de *Yukawa*.

Rappelons qu'un champ scalaire ϕ peut interagir avec deux champs de Dirac ψ_1 et ψ_2 via un couplage simple de la forme suivante :

$$\mathcal{L}_Y = \lambda \bar{\psi}_1 \psi_2 \phi + \text{c.h.} \quad (6.57)$$

Ce terme dans le lagrangien est invariant de Lorentz, et en plus la constante λ est sans unités, ce qui en fait une interaction renormalisable.

Dans le modèle standard, les couplages de ce genre qui ne sont pas interdits par l'invariance de jauge devraient donc être présents. C'est un principe général de la modélisation des interactions fondamentales que si un terme n'est pas interdit par un principe de symétrie, alors il doit être présent dans la théorie. Les couplages de Yukawa possibles doivent cependant obéir aux contraintes suivantes :

1. Ils doivent être invariants lors d'une transformation de jauge $SU(2)$. Comme Φ est un doublet d'isospin faible, l'un des deux champs de Dirac impliqués dans le couplage de Yukawa doit être un doublet conjugué ($\bar{\ell}_L$ ou \bar{q}_L). le deuxième champ de Dirac impliqué est alors nécessairement un singulet droit (e_R , u_R ou d_R).
2. Ils doivent être invariants lors d'une transformation de jauge $U(1)$, ce qui revient à dire que la somme des hypercharges des champs présents doit être nulle. Soulignons cependant qu'un champ conjugué complexe (par exemple $\bar{\ell}_L$ au lieu de ℓ_L) compte pour l'hypercharge opposée, car la transformation de jauge du champ conjugué implique la phase opposée.

Cela laisse trois possibilités :

1. Φ peut être couplé à $\bar{\ell}_L$ et à e_R : $f^{(e)} \bar{\ell}_L e_R \Phi$, où $f^{(e)}$ est une constante sans unités. À cela on doit ajouter le conjugué complexe, car le lagrangien est réel :

$$\mathcal{L}_Y^{(e)} = f^{(e)} \bar{\ell}_L e_R \Phi + f^{(e)} \bar{e}_R \Phi^\dagger \ell_L \quad (6.58)$$

Si on développe cette expression en fonction des composantes des doublets, on trouve

$$f^{(e)}(\bar{\nu}_L, \bar{e}_L) \begin{pmatrix} \phi^+ \\ \phi^0 \end{pmatrix} e_R + \text{c.h.} = f^{(e)}(\phi^+ \bar{\nu}_L e_R + \phi^0 \bar{e}_L e_R) + \text{c.h.} \quad (6.59)$$

En remplaçant le champ de Higgs par sa valeur moyenne ($\phi^+ \rightarrow 0$ et $\phi^0 \rightarrow v/\sqrt{2}$), on trouve

$$\frac{f^{(e)} v}{\sqrt{2}} (\bar{e}_L e_R + \bar{e}_R e_L) = -m_e \bar{e} e \quad (6.60)$$

Ce terme est précisément un terme de masse pour un nouveau champ de Dirac non chiral, formé de e_L et de e_R , soit l'électron, dont la masse est $m_e = -f^{(e)}v/\sqrt{2}$.

2. Φ peut être couplé à \bar{q}_L et à d_R : $f^{(d)}\bar{q}_L d_R \Phi$, où $f^{(e)}$ est une constante sans unités. En suivant le même chemin que pour le terme précédent, on trouve une masse $m_d = -f^{(d)}v/\sqrt{2}$ pour le quark d .
3. La seule autre possibilité est obtenue en considérant le champ conjugué Φ^c , défini par

$$\Phi^c = i\sigma^2 \Phi^* = \begin{pmatrix} \phi^0 \\ -\phi^+ \end{pmatrix} \quad (6.61)$$

Lors d'une transformation de jauge $SU(2)$, le champ Φ^c se transforme exactement comme le champ Φ . En effet, toute matrice de $SU(2)$ peut s'écrire comme

$$U = \cos \theta/2 + i\boldsymbol{\sigma} \cdot \mathbf{n} \sin \theta/2 \quad (6.62)$$

où \mathbf{n} est le vecteur unitaire dans la direction de l'axe de rotation dans l'espace d'isospin faible, et θ l'angle de rotation. Le champ Φ^c se transforme alors de la manière suivante :

$$\Phi^c \rightarrow i\sigma^2 (U\Phi)^* = \sigma^2 U^* \sigma^2 \Phi^c \quad (6.63)$$

où nous avons utilisé le fait que $(\sigma^2)^2 = 1$. D'un autre côté,

$$\sigma^2 U^* \sigma^2 = \cos \theta/2 - i(\sigma^2 \boldsymbol{\sigma}^* \sigma^2) \cdot \mathbf{n} \sin \theta/2 \quad (6.64)$$

Comme σ^2 est purement imaginaire et σ^1 et σ^3 sont réels, $\boldsymbol{\sigma}^* = (\sigma^1, -\sigma^2, \sigma^3)$. D'autre part, σ^2 anticommute avec σ^1 et σ^3 , de sorte que $\sigma^2 \boldsymbol{\sigma}^* \sigma^2 = -\boldsymbol{\sigma}$. On retrouve donc

$$\sigma^2 U^* \sigma^2 = U \quad (6.65)$$

ce qui démontre que Φ^c se transforme comme Φ . Par contre, lors d'une transformation de jauge $U(1)$, Φ^c se transforme avec la phase opposée à celle de Φ , en raison du conjugué complexe. Autrement dit, le passage de Φ à Φ^c équivaut à intervertir les particules et les antiparticules décrites par le champ Φ . Ceci nous permet d'écrire la combinaison invariante de jauge suivante :

$$\mathcal{L}_Y^{(u)} = f^{(u)} \bar{q}_L u_R \Phi^c + \text{c.h.} \quad (6.66)$$

et cette expression mène à un terme de masse pour le quark u lorsqu'on remplace Φ par sa valeur moyenne.

Au total, nous avons donc trois couplages de Yukawa possibles :

$$\mathcal{L}_{fH} = f^{(e)} \bar{\ell}_L e_R \Phi + f^{(u)} \bar{q}_L u_R \Phi^c + f^{(d)} \bar{q}_L d_R \Phi + \text{c.h.} \quad (6.67)$$

où les constantes de couplage $f^{(e,u,d)}$ sont des paramètres supplémentaires de la théorie électro-faible. Lorsque le doublet de Higgs a une valeur moyenne non nulle ($0, v/\sqrt{2}$) dans le vide, ces couplages, en première approximation, deviennent des termes de masse pour les fermions :

$$\mathcal{L}_{fH} \rightarrow f^{(e)} \frac{v}{\sqrt{2}} \bar{\ell}_L e_R + f^{(u)} \frac{v}{\sqrt{2}} \bar{u}_L u_R + f^{(d)} \frac{v}{\sqrt{2}} \bar{d}_L d_R + \text{c.h.} \quad (6.68)$$

et les masses des fermions sont simplement

$$m_e = |f^{(e)}| \frac{v}{\sqrt{2}} \quad m_u = |f^{(u)}| \frac{v}{\sqrt{2}} \quad m_d = |f^{(d)}| \frac{v}{\sqrt{2}} \quad (6.69)$$

ou encore, en fonction de la masse du W ,

$$m_e = \frac{1}{g} \sqrt{2} |f^{(e)}| m_W \quad m_u = \frac{1}{g} \sqrt{2} |f^{(u)}| m_W \quad m_d = \frac{1}{g} \sqrt{2} |f^{(d)}| m_W \quad (6.70)$$

Ainsi, tous les fermions – sauf les neutrinos, qui n'ont pas, dans ce modèle, de composante droite – acquièrent une masse par brisure spontanée de la symétrie de jauge.

Notons ici l'importance philosophique de cette façon de comprendre les masses des particules matérielles. Un terme de masse est par nature *super-renormalisable*, ce qui signifie qu'il est très difficile de comprendre, à partir d'une théorie définie à une échelle fondamentale comme la longueur de Planck, pourquoi les masses de toutes les particules ne seraient pas de l'ordre de la masse de Planck ($\sim 10^{19}$ GeV). Pour cette raison, il est philosophiquement désirable de ne pas avoir de termes de masse dans la théorie des champs de départ. C'est le cas du modèle standard. Les fermions gauches y sont indépendants des fermions droits et se transforment de manière différente sous $SU(2)$, ce qui entraîne qu'on ne peut pas écrire un terme de masse qui soit invariant de jauge. Cependant, les couplages de Yukawa sont permis et, comme ils sont renormalisables, on peut comprendre pourquoi la valeur des constantes f pourrait être de l'ordre de l'unité, car ces couplages ne subissent que des renormalisations logarithmiques en fonction de l'échelle d'énergie. Par le mécanisme de Higgs, ces constantes f , multipliées par la valeur $v/\sqrt{2}$, deviennent des masses dans une théorie effective aux basses énergies (la constante v a les unités d'une masse, car c'est la valeur d'un champ de boson Φ , qui a les unités de la masse). Le fait que des masses puissent exister sans avoir une valeur de l'ordre de la masse de Planck trouve donc une explication naturelle. Cela est l'un des avantages conceptuels majeurs du modèle standard et du mécanisme de brisure spontanée de la symétrie.

Évidemment, il reste un paramètre du modèle standard qui a les unités d'une masse, soit μ dans le potentiel de Higgs (6.20). Ce paramètre, avec λ , détermine la valeur du minimum $v/\sqrt{2}$, qui lui aussi a les unités d'une masse. C'est donc sur cet unique paramètre que repose l'odieux d'être super-renormalisable et plusieurs théories ont été formulées afin de le faire disparaître. L'une de ces idées, appelée « technicolor », suppose que le champ de Higgs n'est pas fondamental, mais est plutôt formé d'une paire de fermions plus fondamentaux, liés par une interaction de jauge supplémentaire. Jusqu'ici rien ne peut confirmer ou infirmer expérimentalement la validité d'une telle hypothèse.

6.D.2 Réplication des familles et mélange des quarks

Une subtilité de la théorie électrofaible apparaît lorsqu'on l'applique non seulement à une famille de particules élémentaires, mais aux trois familles. Rappelons l'agencement des particules élémentaires connues dans ces trois familles, en énumérant tous les doublets d'isospin faible et en introduisant une numérotation des doublets par famille comme suit :

$$\ell_{1L} = \begin{pmatrix} \nu_e \\ e \end{pmatrix}_L \quad \ell_{2L} = \begin{pmatrix} \nu_\mu \\ \mu \end{pmatrix}_L \quad \ell_{3L} = \begin{pmatrix} \nu_\tau \\ \tau \end{pmatrix}_L \quad (6.71)$$

$$q_{1L} = \begin{pmatrix} u \\ d \end{pmatrix}_L \quad q_{2L} = \begin{pmatrix} c \\ s \end{pmatrix}_L \quad q_{3L} = \begin{pmatrix} t \\ b \end{pmatrix}_L \quad (6.72)$$

Utilisons aussi une notation indicielle pour les membres des doublets :

$$\ell_a = \begin{pmatrix} \nu_a \\ e_a \end{pmatrix}_L \quad \text{et} \quad q_a = \begin{pmatrix} p_a \\ n_a \end{pmatrix}_L \quad (6.73)$$

de sorte que

$$e_1 = e \quad e_2 = \mu \quad e_3 = \tau \quad p_1 = u \quad p_2 = c \quad p_3 = t \quad n_1 = d \quad n_2 = s \quad n_3 = b \quad (6.74)$$

Chaque famille (c.-à-d. chaque colonne du tableau des fermions élémentaires) est couplée aux champs de jauge exactement de la même manière. Cependant, les couplages de Yukawa peuvent être plus généraux, et mélanger des fermions de familles différentes, comme suit :

$$\mathcal{L}_{fH} = f_{ab}^{(e)} \bar{\ell}_{a,L} e_{b,R} \Phi + f_{ab}^{(p)} \bar{q}_{a,L} p_{b,R} \Phi^c + f_{ab}^{(n)} \bar{q}_{a,L} n_{b,R} \Phi + \text{c.h.} \quad (6.75)$$

Les termes de masse correspondants sont

$$\mathcal{L}_{fH} \rightarrow \frac{v}{\sqrt{2}} \left\{ f_{ab}^{(e)} \bar{e}_{a,L} e_{b,R} + f_{ab}^{(p)} \bar{p}_{a,L} p_{b,R} + f_{ab}^{(n)} \bar{n}_{a,L} n_{b,R} \right\} + \text{c.h.} \quad (6.76)$$

Nous avons donc affaire à des matrices de masse

$$M_{ab}^{(e,p,n)} = -\frac{v}{\sqrt{2}} f_{ab}^{(e,p,n)} \quad (6.77)$$

qu'il faut diagonaliser afin d'identifier les états propres du hamiltonien sans interactions.

Transformation biunitaire Cette diagonalisation peut se faire à l'aide d'une transformation *bi-unitaire* (ou décomposition en valeurs singulières, DVS) :

$$M = S M_d T^\dagger \quad (6.78)$$

où la matrice M_d est diagonale et les matrices S et T sont unitaires. On montre relativement facilement que toute matrice M peut être diagonalisée de cette façon : la matrice MM^\dagger étant hermitienne, elle peut être diagonalisée par une matrice unitaire U_1 : $U_1 M M^\dagger U_1^\dagger = M_D^2$, où M_D est une matrice diagonale. Le fait que M_D^2 soit définie positive (ses éléments sont tous positifs) provient de ce que MM^\dagger est aussi définie positive ; ceci parce que $x^\dagger M M^\dagger x = |M^\dagger x|^2$ est positif pour tout vecteur x . Ensuite, définissons la matrice $U_2 = M_D^{-1} U_1 M$. Cette matrice est unitaire : on vérifie aisément que $U_2^\dagger U_2 = 1$. Par contre, on constate immédiatement que $M = U_1^\dagger M_D U_2$, ce qui prouve notre assertion initiale.

La diagonalisation est nécessaire afin de pouvoir interpréter les interactions en fonction de particules bien définies. Il s'agit ici de l'équivalent des modes normaux dans un système d'oscillateurs. On doit d'abord identifier les modes normaux avant de traiter les effets non linéaires dans un tel système, c'est-à-dire l'équivalent ici des interactions.

États propres de jauge et états propres de la masse Effectuons un changement de notation : tous les champs considérés précédemment seront affublés d'un prime (') et seront qualifiés d'« états propres de jauge », en ce sens que les courants chargés et neutres sont diagonaux en fonction de ces champs. Les termes de masses, cependant, ne le sont pas, et ont la structure $\bar{\psi}'_L M \psi'_R$, où ψ' désigne e' , p' ou n' . Après diagonalisation, on trouve

$$\bar{\psi}'_L M \psi'_R = \bar{\psi}'_L S M_d T^\dagger \psi'_R = \bar{\psi}_L M_d \psi_R \quad (6.79)$$

où on a défini les champs diagonaux (états propres de la masse) $\psi_L = S^\dagger \psi'_L$ et $\psi_R = T^\dagger \psi'_R$. En inversant ces relations, on trouve

$$\psi'_L = S \psi_L \quad \text{et} \quad \psi'_R = T \psi_R \quad \text{ou, plus spécifiquement,} \quad \begin{cases} e'_L = S^{(e)} e_L & e'_R = T^{(e)} e_R \\ p'_L = S^{(p)} p_L & p'_R = T^{(p)} p_R \\ n'_L = S^{(n)} n_L & n'_R = T^{(n)} n_R \end{cases} \quad (6.80)$$

Ceci implique que les courants chargés et neutres ne sont pas diagonaux dans les nouveaux champs e , p et n et donc que l'interaction faible peut effectuer des transitions entre les différentes familles. Voyons comment cela fonctionne : la contribution des quarks au courant chargé $J^{+\mu}$ devient, en fonctions des états propres de la masse,

$$J_{\text{quarks}}^{+\mu} = \bar{q}'_L \gamma^\mu \sigma^+ q'_L = \bar{p}'_L \gamma^\mu n'_L = \bar{p}_L S^{(p)\dagger} \gamma^\mu S^{(n)} n_L = \bar{p}_L \gamma^\mu U n_L \quad (6.81)$$

où la matrice $U = S^{(p)\dagger} S^{(n)}$ est unitaire, et caractérise un mélange des quarks (d , s , b) se couplant au quark u dans le courant chargé.

Ceci signifie, par exemple, que dans les diagrammes de la fig. 6.6, le quark d devrait être remplacé par la combinaison $U_{11}d + U_{12}s + U_{13}b$, et donc que des processus sont possibles par lesquels un quark s se désintègre en quark u et en leptons. La matrice U porte le nom de matrice de Cabibbo-Kobayashi-Maskawa (matrice CKM), et ses éléments forment aussi des paramètres du modèle standard (en fait, seuls quatre paramètres sont vraiment indépendants dans la spécification de cette matrice). Plusieurs paramétrisations de cette matrice ont été proposées. La plus utilisée est la suivante :

$$U = \begin{pmatrix} c_{12}c_{13} & s_{12}c_{13} & s_{13}e^{-i\delta_{13}} \\ -s_{12}c_{23} - c_{12}s_{23}s_{13}e^{i\delta_{13}} & c_{12}c_{23} - s_{12}s_{23}s_{13}e^{i\delta_{13}} & s_{23}c_{13} \\ s_{12}s_{23} - c_{12}c_{23}s_{13}e^{i\delta_{13}} & -c_{12}s_{23} - s_{12}c_{23}s_{13}e^{i\delta_{13}} & c_{23}c_{13} \end{pmatrix} \quad (6.82)$$

où $c_{ij} = \cos \theta_{ij}$ et $s_{ij} = \sin \theta_{ij}$. Les quatre paramètres utilisés ici sont θ_{12} , θ_{13} , θ_{23} et δ_{13} . Les valeurs absolues des éléments de la matrice CKM sont estimées à

$$|U_{ij}| = \begin{pmatrix} 0,97383 & 0,2272 & 0,00396 \\ 0,2271 & 0,97296 & 0,04221 \\ 0,00814 & 0,04161 & 0,999100 \end{pmatrix} \quad (6.83)$$

Angle de Cabibbo Le cas de deux familles est plus simple, et a été étudié plus tôt. Dans ce cas, la matrice CKM est 2×2 et ne comporte comme paramètre indépendant qu'un

angle θ_c appelé *angle de Cabibbo* :

$$U = \begin{pmatrix} \cos \theta_c & \sin \theta_c \\ -\sin \theta_c & \cos \theta_c \end{pmatrix} \quad (6.84)$$

ce qui donne le courant chargé hadronique suivant :

$$\begin{aligned} J^{+\mu} &= \bar{p}_L \gamma^\mu U n_L = (\bar{u}_L \quad \bar{c}_L) \gamma^\mu \begin{pmatrix} \cos \theta_c & \sin \theta_c \\ -\sin \theta_c & \cos \theta_c \end{pmatrix} \begin{pmatrix} d_L \\ s_L \end{pmatrix} \\ &= \cos \theta_c \bar{u}_L \gamma^\mu d_L + \sin \theta_c \bar{u}_L \gamma^\mu s_L - \sin \theta_c \bar{c}_L \gamma^\mu d_L + \cos \theta_c \bar{c}_L \gamma^\mu s_L \end{aligned} \quad (6.85)$$

De cette manière, on conclut que l'amplitude du processus de désintégration $s \rightarrow ue \nu$ est $\tan \theta_c$ fois l'amplitude correspondante pour la désintégration du quark d . La valeur observée de l'angle de Cabibbo est $\theta_c = 0,222$. Mais gardons à l'esprit que la matrice CKM est en réalité 3×3 et que d'autres angles et phases interviennent. Mais le mélange entre la première et la deuxième famille est le plus important.

6.D.3 Cas des leptons et des courants neutres

Les neutrinos n'ayant pas de masse (en raison de l'absence des neutrinos droits ν_R), la contribution des leptons au courant chargé s'écrit plutôt

$$J_{\text{lep}}^{+\mu} = \bar{\ell}'_L \gamma^\mu \sigma^+ \ell'_L = \bar{\nu}'_L \gamma^\mu e'_L = \bar{\nu}'_L \gamma^\mu S^{(e)} e_L \quad (6.86)$$

Mais les neutrinos n'ayant pas de masse, les trois espèces sont dégénérées. Donc on peut sans peine définir $\nu = S^{(e)\dagger} \nu'$ sans affecter les masses (nulles) des neutrinos : ν' et ν représentent tous les deux des états propres de la masse. Le courant chargé leptonique reste donc diagonal en fonction des familles de leptons :

$$J_{\text{lep}}^{+\mu} = \bar{\nu}_L \gamma^\mu e_L \quad (6.87)$$

Voir cependant la question 6.2.

Courants neutres Les courants neutres ne sont pas non plus affectés par le changement de base dans l'espace des familles, c'est-à-dire que leur forme demeure la même qu'avant la transformation, simplement parce qu'ils sont diagonaux. En effet,

$$\begin{aligned} J^{3\mu} &= \frac{1}{2} \bar{\ell}'_L \gamma^\mu \sigma^3 \ell'_L + \frac{1}{2} \bar{q}'_L \gamma^\mu \sigma^3 q'_L \\ &= \frac{1}{2} \bar{\nu}'_L \gamma^\mu \nu'_L - \frac{1}{2} \bar{e}'_L \gamma^\mu e'_L + \frac{1}{2} \bar{p}'_L \gamma^\mu p'_L - \frac{1}{2} \bar{n}'_L \gamma^\mu n'_L \\ &= \frac{1}{2} \bar{\nu}_L S^{(e)\dagger} \gamma^\mu S^{(e)} \nu_L - \frac{1}{2} \bar{e}_L S^{(e)\dagger} \gamma^\mu S^{(e)} e_L + \frac{1}{2} \bar{p}_L S^{(p)\dagger} \gamma^\mu S^{(p)} p_L - \frac{1}{2} \bar{n}_L S^{(n)\dagger} \gamma^\mu S^{(n)} n_L \\ &= \frac{1}{2} \bar{\nu}_L \gamma^\mu \nu_L - \frac{1}{2} \bar{e}_L \gamma^\mu e_L + \frac{1}{2} \bar{p}_L \gamma^\mu p_L - \frac{1}{2} \bar{n}_L \gamma^\mu n_L \end{aligned} \quad (6.88)$$

car les matrices $S(e, p, n)$ sont diagonales. Le résultat est donc encore diagonal dans l'espace des familles et les courants neutres (c'est-à-dire le Z) ne peuvent pas induire de changement de famille. Tout phénomène découlant d'un courant neutre modifiant la famille (angl. *flavor-changing neutral current*) serait au-delà du modèle standard et n'a jamais été observé. Il en va bien sûr de même du courant électromagnétique.

Résumé : Lagrangien du modèle standard avec trois familles de fermions

On néglige ici les fluctuations du champ de Higgs : seule la valeur moyenne est prise en compte et se retrouve cachée dans les masses.

$$\begin{aligned}
 \mathcal{L} = & -\frac{1}{4}(\partial_\mu A_\nu - \partial_\nu A_\mu)(\partial^\mu A^\nu - \partial^\nu A^\mu) - \frac{1}{4}(\partial_\mu Z_\nu - \partial_\nu Z_\mu)(\partial^\mu Z^\nu - \partial^\nu Z^\mu) && \text{termes de Maxwell} \\
 & -\frac{1}{2}(\partial_\mu W_\nu - \partial_\nu W_\mu)^*(\partial^\mu W^\nu - \partial^\nu W^\mu) \\
 & + m_W^2 W^{\mu*} W_\mu + \frac{1}{2} m_Z^2 Z^\mu Z_\mu && \text{masses des champs de jauge} \\
 & + \frac{g^2}{4}(W_\mu^* W_\nu - W_\nu^* W_\mu)(W^{\mu*} W^\nu - W^{\nu*} W^\mu) && \text{interactions entre champs de jauge} \\
 & - \frac{g^2}{2} \sin^2 \theta_W [2A_\mu A^\mu W_\nu^* W^\nu - A^\mu A^\nu (W_\mu^* W_\nu + W_\nu^* W_\mu)] \\
 & - \frac{g^2}{2} \cos^2 \theta_W [2Z_\mu Z^\mu W_\nu^* W^\nu - Z^\mu Z^\nu (W_\mu^* W_\nu + W_\nu^* W_\mu)] \\
 & - \frac{g^2}{2} \sin \theta_W \cos \theta_W [4A_\mu Z^\mu W_\nu^* W^\nu - (A^\mu Z^\nu + A^\nu Z^\mu)(W_\mu^* W_\nu + W_\nu^* W_\mu)] \\
 & + g \sin \theta_W \text{Im}[(\partial_\mu W_\nu - \partial_\nu W_\mu)^*(A^\mu W^\nu - A^\nu W^\mu) + (\partial_\mu A_\nu - \partial_\nu A_\mu)(W^{\mu*} W^\nu - W^{\nu*} W^\mu)] \\
 & + g \cos \theta_W \text{Im}[(\partial_\mu W_\nu - \partial_\nu W_\mu)^*(Z^\mu W^\nu - Z^\nu W^\mu) + (\partial_\mu Z_\nu - \partial_\nu Z_\mu)(W^{\mu*} W^\nu - W^{\nu*} W^\mu)] \\
 & + i \bar{\nu}_L^a \gamma^\mu \partial_\mu \nu_L^a + i \bar{e}^a \gamma^\mu \partial_\mu e^a + i \bar{p}^a \gamma^\mu \partial_\mu p^a + i \bar{n}^a \gamma^\mu \partial_\mu n^a && \text{termes cinétiques} \\
 & - m^{(e,a)} \bar{e}^a e^a - m^{(p,a)} \bar{p}^a p^a - m^{(n,a)} \bar{n}^a n^a && \text{masses des fermions} \\
 & - e A \cdot J_{\text{em}} - \frac{g}{\cos \theta_W} Z \cdot J_0 - \frac{g}{\sqrt{2}} (J \cdot W + J^* \cdot W^*) && \text{interactions boson-fermion}
 \end{aligned}$$

où les courants sont donnés par

$$\begin{aligned}
 J_{\text{em}}^\mu &= -\bar{e}^a \gamma^\mu e^a + \frac{2}{3} \bar{p}^a \gamma^\mu p^a - \frac{1}{3} \bar{n}^a \gamma^\mu n^a && \text{courant électromagnétique} \\
 J^\mu &= \bar{\nu}_L^a \gamma^\mu e_L^a + U_{ab} \bar{p}_L^a \gamma^\mu n_L^b && U : \text{matrice CKM} \\
 &&& \text{courant chargé} \\
 J^{3\mu} &= \frac{1}{2} (\bar{\nu}_L^a \gamma^\mu \nu_L^a - \bar{e}_L^a \gamma^\mu e_L^a + \bar{p}_L^a \gamma^\mu p_L^a - \bar{n}_L^a \gamma^\mu n_L^a) \\
 J_0^\mu &= J^{3\mu} - \sin^2 \theta_W J_{\text{em}}^\mu && \text{courant neutre}
 \end{aligned}$$

On somme sur les indices répétés. Les champs de fermions représentent les particules suivantes, avec les mêmes symboles :

$\nu^1 = \nu_e$	$\nu^2 = \nu_\mu$	$\nu^3 = \nu_\tau$	neutrinos
$e^1 = e$	$e^2 = \mu$	$e^3 = \tau$	leptons chargés $Q = -1$
$p^1 = u$	$p^2 = c$	$p^3 = t$	quarks $Q = \frac{2}{3}$
$n^1 = d$	$n^2 = s$	$n^3 = b$	quarks $Q = -\frac{1}{3}$

Constantes de couplage et angle de Weinberg :

$$\sin^2 \theta_W = 0,23120(15) \quad \frac{g^2}{4\pi} \approx (31,683)^{-1} \quad e = g \sin \theta_W \quad m_W = 80,398 \text{ GeV} \quad m_Z = 91,19 \text{ GeV}$$

E Physique des neutrinos

6.E.1 Détection des neutrinos

La première « détection » des neutrinos a eu lieu en 1955 et a valu un prix Nobel à F. REINES. La réaction en jeu était la suivante :

$$p + \bar{\nu}_e \rightarrow n + e^+ \quad (6.89)$$

Les antineutrinos provenaient de la désintégration de neutrons produits par un réacteur nucléaire. Comme les neutrinos traversent aisément tout matériau, il est possible d'aménager un détecteur à l'extérieur du réacteur sans qu'aucun neutron en provenance du même réacteur ne puisse y parvenir, en raison de l'enveloppe du réacteur. Les antineutrinos frappent alors une cible contenant naturellement une forte densité de protons. Le positron produit par la réaction s'annihile avec un électron pour produire deux photons d'énergies pratiquement égales. Le neutron produit, lui, est capturé par du Cadmium présent dans la cible et cette capture est suivie d'une émission gamma de 8 MeV par le noyau de Cadmium excité. La détection de la paire de photons et du gamma à 8 MeV à quelques microsecondes d'intervalle est un signe certain que la réaction s'est produite.

Les antineutrinos ne sont donc pas détectés directement : ce sont les produits d'une réaction les impliquant qui sont détectés. Mais cela n'a rien d'inhabituel, car en y pensant bien, toutes les particules sauf les photons visibles sont détectées de cette manière. La seule différence réside dans la chaîne de réactions nécessaire, qui est légèrement plus complexe pour les neutrinos que pour des neutrons, par exemple.

L'un des détecteurs de neutrinos en service jusqu'à récemment est situé à l'observatoire de neutrinos de Sudbury (SNO). Ce détecteur était particulier en ce qu'il était basé sur des réactions impliquant le deutéron. Comme le Canada produit de grandes quantités d'eau lourde en raison de sa filière nucléaire particulière, il constitue un endroit privilégié pour accueillir un tel détecteur. Le SNO était enfoui à 2 000m sous la terre, dans une mine de Nickel. Il était constitué d'une cavité sphérique de 12m de diamètre remplie d'eau lourde (1 000 tonnes) et bordée de 9 600 photomultiplicateurs. La grande profondeur du détecteur était nécessaire afin de s'assurer qu'il soit à l'abri des rayons cosmiques et que seuls les neutrinos puissent y parvenir. Les réactions observables dans ce détecteur étaient les suivantes :

$$\begin{aligned} \nu_e + d &\rightarrow p + p + e \\ \nu + d &\rightarrow p + n + \nu \\ \nu + e &\rightarrow \nu + e \end{aligned} \quad (6.90)$$

La première réaction ne peut faire intervenir que des neutrinos électroniques, car les autres saveurs de neutrinos, s'ils proviennent des réactions solaires, n'auraient pas l'énergie suffisante pour créer un muon ou un tau. Par contre, les deux autres réactions sont possibles avec n'importe laquelle saveur de neutrino. La deuxième réaction est une dissociation du deutéron induite par un neutrino, et la troisième une simple diffusion élastique avec un électron. Les particules chargées parmi les produits sont détectées par rayonnement Tchénkov, c'est-à-dire qu'elles vont plus vite que la

lumière dans l'eau et donc qu'elles donnent naissance à un front d'onde conique dont l'angle est directement relié à la vitesse de la particule.

6.E.2 Faisceaux de neutrinos

Il est possible d'aménager des faisceaux de neutrinos, même si on ne peut pas les contrôler directement. Le principe de base est illustré à la figure 6.7. Un faisceau de protons de haute énergie est envoyé sur une cible, générant ainsi des pions et des kaons secondaires. Ces secondaires sont ensuite filtrés dans un collimateur qui, à l'aide d'aimants dipolaires et quadripolaires, produit un faisceau de pions et de kaons beaucoup plus homogène en quantité de mouvement. Ce faisceau épuré parcourt ensuite un tunnel de désintégration dans lequel la plus grande partie des hadrons s'est désintégrée en neutrinos muoniques et en muons. La prochaine étape du parcours est un tampon d'acier qui élimine les muons du faisceau, ne laissant que les neutrinos muoniques. Un filtre de terre, beaucoup plus long, suffit à éliminer toutes les particules résiduelles. Il reste un faisceau de neutrinos, assez bien résolu en énergie, qui progresse vers le détecteur.

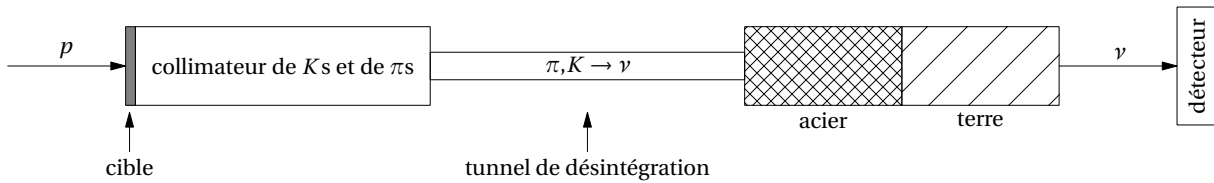


FIGURE 6.7

Génération d'un faisceau de neutrinos. Les différentes composantes ne sont pas à l'échelle. En particulier, plusieurs centaines de kilomètres peuvent séparer le détecteur de neutrinos du reste.

C'est grâce à un tel faisceau artificiel de neutrinos que les premiers signes de l'existence du courant neutre ont été observés, au CERN, en 1973. La réaction impliquée est la diffusion entre un électron et un neutrino muonique :

$$\nu_\mu + e \rightarrow \nu_\mu + e \quad (6.91)$$

qui ne peut être médiée que par un Z , en vertu du diagramme suivant :



6.E.3 Oscillations de neutrinos

L'un des paradoxes les plus connus entourant le modèle standard est le phénomène de l'oscillation des neutrinos. Le problème est le suivant : à la fin des années 1960, R. Davis et son équipe construisirent un détecteur dont l'objectif était de mesurer le flux de neutrinos électroniques en provenance

du Soleil. Le flux de neutrinos solaire est prédit avec une bonne précision par un modèle très solide des réactions nucléaires au coeur du soleil. Le résultat essentiel de cette expérience fut que seulement un tiers des neutrinos attendus était observé. Ce paradoxe fut appelé par la suite le « problème des neutrinos solaires ». Une solution possible à ce paradoxe était une remise en question du modèle solaire, ce qui s'avéra irréaliste en raison de la robustesse de celui-ci. La solution retenue fut plutôt l'existence d'oscillations entre les trois saveurs de neutrinos (ν_e , ν_μ et ν_τ). Cela signifie qu'un neutrino, créé en tant que ν_e , se transforme au cours de sa propagation en ν_μ et en ν_τ (on dit qu'il « oscille » entre les trois saveurs de neutrinos). Ceci est possible si les états propres de la masse ne sont pas les mêmes que les états propres de jauge, comme pour les quarks, et requiert bien évidemment que les neutrinos aient une masse, aussi faible soit-elle.

Dans les années 1990s, grâce à des détecteurs sensibles aux trois espèces de neutrinos, tel le SNO, on a pu confirmer que le flux total de neutrinos est bel et bien trois fois le flux des neutrinos électroniques, ce qui résout de manière satisfaisante le problème des neutrinos solaires.

Reste à savoir d'où vient la masse des neutrinos et quelle forme précise elle prend. Voir à cet effet le problème 6.2.

6.E.4 Désintégration du pion chargé

Le pion chargé (π^\pm) se désintègre, via émission d'un W virtuel, en un lepton et son neutrino (voir fig. 6.8). Étant donnée la masse du pion (139 MeV), les seules possibilités sont

$$\pi^- \rightarrow \mu + \bar{\nu}_\mu \quad \text{et} \quad \pi^- \rightarrow e + \bar{\nu}_e \quad (6.93)$$

Or, il se trouve que le rapport d'embranchement Γ_μ/Γ_e entre ces deux processus est d'environ 8 000. Autrement dit, il est 8 000 fois plus probable que le pion se désintègre en muon qu'en électron. À première vue, cela peut paraître paradoxal, car l'espace des phases favorise le deuxième processus, la masse de l'électron étant très petite par rapport à celle du pion. Mais nous allons voir que c'est justement la petitesse de la masse de l'électron qui en fait un canal de désintégration si improbable pour le pion. Un argument simple est le suivant : plaçons-nous dans le repère du pion. Son spin étant nul, il faut que le moment cinétique total des produits de sa désintégration soit nul lui aussi. L'électron et l'antineutrino s'en vont dans des directions opposées. Comme l'antineutrino est sans masse, son spin est dans la même direction que sa quantité de mouvement (seuls les antineutrinos droits existent). Si l'électron était lui aussi sans masse, son spin serait dans la direction opposée à sa quantité de mouvement, soit la même direction que le spin de l'antineutrino, car seule la composante gauche de l'électron intervient dans l'interaction faible, responsable de la désintégration du pion. Ceci donnerait une projection totale du spin égale à 1 dans la direction de l'antineutrino, en conflit avec la conservation du moment cinétique et donc interdit. Comme la masse de l'électron n'est pas nulle, le processus n'est pas complètement interdit, mais simplement fortement désavantagé par rapport au canal muonique. En fait, l'amplitude de la composante droite de l'électron est de l'ordre de m/E , E étant l'énergie de l'électron et m sa masse, ce qui mène à une probabilité relative de l'ordre de $(m/E)^2$ pour le canal électronique. Comme $E \sim 70\text{MeV}$ et $m \sim 0,5\text{MeV}$, ce rapport est d'environ $140^{-2} \sim 1/20\,000$. En tant compte cependant de l'avantage apporté par le plus grand espace de phases associé au canal électronique, ce rapport est augmenté à $1/8\,000$.

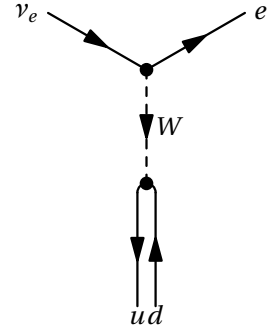

FIGURE 6.8

Diagramme décrivant la désintégration du pion π^- en électron et antineutrino.

Nous allons maintenant procéder à un calcul plus détaillé. Le lagrangien (6.38) est à l'origine de ce processus. Le vertex correspondant doit tenir compte de ce qu'uniquement les fermions gauches interviennent, ce qui mène à un facteur

$$-\frac{ig}{2\sqrt{2}}\gamma^\mu(1-\gamma^5) \quad (6.94)$$

D'autre part, on peut utiliser l'approximation de basse énergie pour le propagateur du W , à savoir

$$-i\frac{g_{\mu\nu}}{m_W^2} \quad (6.95)$$

En numérotant l'antineutrino (2) et l'électron (3), les règles de Feynman mènent à l'amplitude suivante :

$$\mathcal{M} = \frac{g^2}{8m_W^2} [\bar{u}(3)\gamma_\mu(1-\gamma^5)v(2)]F^\mu \quad (6.96)$$

où F^μ est un quadrivecteur représentant la partie « pion » de l'amplitude. Le seul quadrivecteur disponible est la quadri-impulsion p^μ du pion, et donc nous allons supposer que ce facteur a la forme $F^\mu = f_\pi p^\mu$, où f_π est une constante inconnue ayant les unités d'une masse. On montre ensuite (voir problème 6.5) que l'amplitude carrée sommée sur les spins des particules finales est

$$\langle |\mathcal{M}|^2 \rangle = \left(\frac{f_\pi g^2}{8m_W^2} \right)^2 p_\mu p_\nu \text{tr} \{ \gamma^\mu(1-\gamma^5) \not{p}_2 \gamma^\nu(1-\gamma^5)(\not{p}_3 + m) \} \quad (6.97)$$

Le calcul de la trace donne

$$\langle |\mathcal{M}|^2 \rangle = \frac{1}{8} \left(\frac{f_\pi g^2}{m_W^2} \right)^2 \{ 2(p_2 \cdot p)(p_3 \cdot p) - m_\pi^2 p_2 \cdot p_3 \} \quad (6.98)$$

En tenant compte de la loi de conservation $p = p_2 + p_3$, on trouve enfin

$$\langle |\mathcal{M}|^2 \rangle = f_\pi^2 \left(\frac{g}{2m_W} \right)^4 m_e^2 (m_\pi^2 - m_e^2) \quad (6.99)$$

D'après la formule (1.82), le taux de désintégration est ensuite

$$\Gamma_{\pi \rightarrow e} = \frac{f_\pi^2}{\pi} \left(\frac{g}{4m_W} \right)^4 \frac{m_e^2}{m_\pi^3} (m_\pi^2 - m_e^2)^2 \quad (6.100)$$

Le rapport entre le taux du canal électronique et le taux du canal muonique est alors

$$\frac{\Gamma_{\pi \rightarrow e}}{\Gamma_{\pi \rightarrow \mu}} = \frac{m_e^2}{m_\mu^2} \left(\frac{m_\pi^2 - m_e^2}{m_\pi^2 - m_\mu^2} \right)^2 \quad (6.101)$$

En substituant les valeurs des masses, on trouve que ce rapport vaut $1,28 \times 10^{-4}$. L'expérience donne plutôt la valeur $(1,23 \pm 0,02) \times 10^{-4}$, ce qui constitue une excellente confirmation de l'hypothèse faite sur la forme de F^μ . La comparaison avec la mesure de $\Gamma_{\pi \rightarrow \mu}$ donne d'ailleurs la valeur $f_\pi = 93 \text{ MeV}$.

F La symétrie CP et sa brisure

l'interaction faible brise la parité, car seuls les fermions gauches interviennent dans la symétrie de jauge SU(2). Cette violation est maximale, en ce sens qu'elle n'est pas caractérisée par un petit paramètre, mais est aussi grande qu'elle pourrait l'être. Cette brisure de la parité P s'accompagne d'une brisure correspondante de la conjugaison de charge C .

En effet, lorsqu'on applique l'opérateur de conjugaison de charge sur un spineur droit, on obtient un spineur gauche :

$$C\psi_R = i\gamma^2\psi_R^* = \frac{1}{2}i\gamma^2((1+\gamma^5)\psi)^* = \frac{1}{2}(1-\gamma^5)i\gamma^2\psi^* = (\psi_c)_L \quad (6.102)$$

parce que γ^5 anticommute avec γ^2 . En clair, l'opération de conjugaison de charge remplace la composante droite d'une particule par la composante gauche de son antiparticule. Or, dans la théorie électrofaible, on trouve des neutrinos gauches et des antineutrinos gauches, mais pas de neutrinos et d'antineutrinos droits. Donc cette théorie n'est pas invariante par conjugaison de charge.

Notons cependant qu'elle est invariante sous l'action combinée de la parité et de la conjugaison de charge. L'opération de parité échange les composantes droite et gauche, de sorte que l'opération CP change un neutrino droit en un antineutrino droit, et la théorie est bel et bien invariante par cette opération. On dit alors qu'elle conserve CP .

L'une des conséquences de cette conservation de CP est le mélange des kaons neutres, expliqué par Gell-Man et Pais en 1955. Les kaons neutres K^0 et \bar{K}^0 sont les antiparticules l'une de l'autre. Leur contenu en quarks est

$$K^0 = \bar{s}d \quad \bar{K}^0 = s\bar{d} \quad (6.103)$$

Ces particules sont, en quelque sorte, des états propres des interactions fortes, qui les créent sous cette forme lors de réactions hadroniques. Par contre, ces particules ne sont pas des états propres de C , car, symboliquement, $C|K^0\rangle = |\bar{K}^0\rangle$ et $C|\bar{K}^0\rangle = |K^0\rangle$. Ce sont cependant des états propres de P , avec une parité intrinsèque $\eta = -1$. Comme la désintégration de ces particules est gouvernée par l'interaction faible, il est normal de chercher à les décrire par des états propres de CP , qui est conservé par l'interaction faible. On construit donc les combinaisons linéaires suivantes :

$$|K_1\rangle = \frac{1}{\sqrt{2}}(|K^0\rangle - |\bar{K}^0\rangle) \quad |K_2\rangle = \frac{1}{\sqrt{2}}(|K^0\rangle + |\bar{K}^0\rangle) \quad (6.104)$$

qui sont des états propres de CP avec valeurs propres $+1$ et -1 respectivement. Par exemple,

$$CP|K_1\rangle = \frac{1}{\sqrt{2}}(CP|K^0\rangle - CP|\bar{K}^0\rangle) = \frac{1}{\sqrt{2}}\eta(C|K^0\rangle - C|\bar{K}^0\rangle) = \frac{1}{\sqrt{2}}\eta(|\bar{K}^0\rangle - |K^0\rangle) = |K_1\rangle \quad (6.105)$$

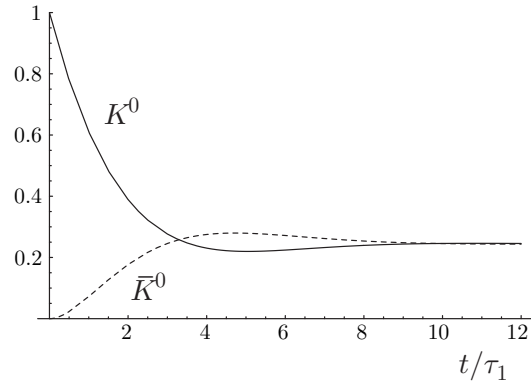
Les deux états K_1 et K_2 ne se désintègrent pas de la même façon. K_1 peut se désintégrer en deux pions : $K_1 \rightarrow \pi^+\pi^-$ ou $K_1 \rightarrow \pi^0\pi^0$. En effet, plaçons-nous dans le référentiel du kaon ; l'état initial et l'état final ont un spin nul et le moment cinétique orbital des deux pions est donc nul. La parité de la fonction d'onde des deux pions est donc $+1$. Lorsqu'on applique une inversion de l'espace, les deux pions, qui s'en vont dans des directions opposées, sont échangés, et ils sont échangés de nouveau lorsqu'on applique la conjugaison de charge, car ils sont l'antiparticule l'un de l'autre (le π^0 est, quant à lui, sa propre antiparticule, ce qui revient au même). Donc l'état final est invariant par CP , c'est-à-dire qu'il est un état propre de CP avec valeur propre $+1$. La désintégration est donc permise, en vertu de la conservation de CP dans l'interaction faible. Par contre, K_2 ayant $CP = -1$ – c'est-à-dire, une valeur propre de CP égale à -1 – ne peut pas se désintégrer en deux pions, mais il peut se désintégrer en trois pions : $K_2 \rightarrow \pi\pi\pi$. Comme l'énergie cinétique disponible dans l'état final est moindre dans ce cas, l'espace des phases est plus restreint et le taux de désintégration est plus faible, menant à des vies moyennes différentes :

$$\tau_1 = 0,9 \times 10^{-10} \text{ s} \quad \tau_2 = 0,5 \times 10^{-7} \text{ s} \quad (6.106)$$

Le K_2 vit en moyenne 550 fois plus longtemps que le K_1 .

FIGURE 6.9

Évolution dans le temps de la probabilité de trouver un kaon K^0 et un antikaon \bar{K}^0 dans un faisceau ne contenant initialement que des K^0 .



Nous pouvons caractériser la propagation de ces particules en spécifiant leur masse et leur vie moyenne inverse $\Gamma = \tau^{-1}$. En fonction du temps, l'état $|K_1\rangle$ évolue de la manière suivante :

$$|K_1(t)\rangle = a_1(t)|K_1\rangle \quad \text{où} \quad a_1(t) = a_1(0)e^{-im_1 t - \Gamma_1 t/2} \quad (6.107)$$

et pareillement pour K_2 . Les conditions initiales font que seuls les K_0 sont produits à $t = 0$. En effet, la production associée de quarks s et \bar{s} fait que l'état $\bar{s}d$ est formé, mais pas l'état $s\bar{d}$, simplement parce qu'il n'y a pas de quark \bar{d} dans la cible frappée par le faisceau de l'accélérateur. Le \bar{s} se combine à d (ou à u) pour former un kaon, et le s au reste des quarks pour former un baryon. Comme

$$K^0 = \frac{1}{\sqrt{2}}(|K_1\rangle + |K_2\rangle) \quad \bar{K}^0 = \frac{1}{\sqrt{2}}(|K_1\rangle - |K_2\rangle), \quad (6.108)$$

les conditions initiales sont $a_1(0) = a_2(0) = 1/\sqrt{2}$, et la probabilité de trouver un K^0 ou un \bar{K}^0 au temps t dans le faisceau de kaons ainsi créé est

$$\begin{aligned} I(K^0) &= \frac{1}{2}|a_1(t) + a_2(t)|^2 = \frac{1}{4}(e^{-\Gamma_1 t} + e^{-\Gamma_2 t} + 2e^{-(\Gamma_1 + \Gamma_2)t/2} \cos(\Delta m t)) \\ I(\bar{K}^0) &= \frac{1}{2}|a_1(t) - a_2(t)|^2 = \frac{1}{4}(e^{-\Gamma_1 t} + e^{-\Gamma_2 t} - 2e^{-(\Gamma_1 + \Gamma_2)t/2} \cos(\Delta m t)) \end{aligned} \quad (6.109)$$

où $\Delta m = |m_1 - m_2|$. Ce comportement est illustré à la fig. 6.9. On a mesuré que $\Delta m \tau_1 = 0,477 \pm 0,002$. Ceci équivaut à une différence de masse $\Delta m = 3,49 \times 10^{-12}$ MeV.

Nous avons vu que les K^0 sont produits par impact d'un faisceau de protons sur une cible, alors que les \bar{K}^0 ne peuvent être produits de cette façon, car ils contiennent un quark \bar{d} qui n'existe pas dans la cible. Ce quark est généré par l'interaction faible au cours de la propagation du kaon. Lorsque le faisceau de kaons frappe une cible, la situation est inversée : l'antikaon \bar{K}^0 peut réagir facilement avec la matière de la cible, car le quark \bar{d} qu'il contient peut facilement s'annihiler avec les quarks d de la cible, alors que le quark s du kaon ne peut pas s'annihiler avec un anti- s , absent de la cible. Donc les anti-kaons réagiront facilement avec la cible pour produire d'autre chose alors que les kaons seront beaucoup moins affectés par la cible et pourront la traverser intacte (ce sont toutes les deux des particules neutres et donc seules les interactions fortes les affectent d'une manière importante dans la cible). Ce phénomène est appelé régénération, même si en réalité les kaons ne sont pas régénérés : ce sont les anti-kaons qui sont éliminés du faisceau qui traverse la cible. En mesurant le flux de kaons sortant de la deuxième cible, en fonction de la distance entre la première cible (qui produit les kaons en premier lieu) et la deuxième, on arrive à mesurer l'amplitude $|a_1(t) + a_2(t)|^2$ et à confirmer les prédictions de la fig. 6.9.

La physique des kaons est une application particulièrement frappante de la mécanique quantique d'un système à deux niveaux, et illustre très bien la notion de particule en mécanique quantique, avec les superpositions possibles qu'elle implique.

6.F.1 Brisure de CP

Les kaons neutres existent donc en deux espèces, qu'on a appelé K_1 et K_2 , mais qu'on devrait de manière plus prudente appeler K_S^0 et K_L^0 (S et L pour *short* et *long*) désignant respectivement les états à vie moyenne courte et longue. Si la symétrie CP est exacte, alors K_S coïncide avec K_1 et K_L avec K_2 .

En 1964, les équipes de Cronin et Fitch ont montré que le kaon K_L , qui en principe ne devrait pas se désintégrer en deux pions en raison de la conservation de CP , parvient quand même à se désintégrer dans ce canal, avec un rapport d'embranchement toutefois assez petit :

$$\frac{K_L^0 \rightarrow \pi^+ \pi^-}{K_L^0 \rightarrow \text{modes chargés}} = 0,002 \quad (6.110)$$

Donc K_L^0 , un état propre de l'interaction faible, n'est pas un état propre parfait de CP ; il y a une petite déviation. Cette révélation créa une onde de choc dans le monde de la physique des particules, comparable à celle qu'avait créée la découverte de la brisure de la parité en 1957.

Un théorème très robuste de la physique théorique des particules stipule que la combinaison des opérations C, P, et T doit vraiment laisser l'univers invariant. En clair, selon ce *théorème CPT*, notre univers est entièrement équivalent à un autre univers obtenu en inversant l'espace, en inversant le cours du temps et en remplaçant la matière par l'antimatière. Il découle de ce théorème, par exemple, que la masse des particules doit être rigoureusement identique à celle des antiparticules. Pratiquement personne ne met en doute la conservation de CPT.

La violation de *CP* entraîne donc que la symétrie T est violée également, c'est-à-dire que les lois microscopiques de la physique ne sont pas invariantes par inversion du temps. Cette asymétrie ne se manifeste que dans l'interaction faible, et de manière assez petite, mais elle existe tout de même.

L'importance de la brisure de *CP* vient de la cosmologie. Notre univers comporte, heureusement, un excédent de matière sur l'antimatière. Le nombre de nucléons dans l'univers est environ un milliard de fois plus petit que le nombre de photons. Si on suppose que ces photons sont venus de l'annihilation des quarks avec les antiquarks (et des électrons avec les positrons), on doit supposer qu'à l'époque où le rayonnement était en équilibre avec la matière (c'est-à-dire qu'il y avait autant de créations de paires particule-antiparticule que d'annihilations), l'excédent de quarks sur les antiquarks devrait être de l'ordre d'un milliardième :

$$\frac{n_q - n_{\bar{q}}}{n_q + n_{\bar{q}}} \sim 10^{-9} \quad (6.111)$$

Cette petite différence a permis à la matière de survivre ; autrement, l'univers ne contiendrait que des photons. En 1967, le physicien russe A. Sakharov – plus connu en tant que père de la bombe H soviétique et plus tard comme le plus célèbre des dissidents russes – a énoncé trois conditions qui, ensemble, peuvent expliquer l'excédent de matière sur l'antimatière dans notre univers :

1. Une brisure de l'inversion du temps dans les lois fondamentales.
2. Une phase dans l'évolution de l'Univers, peu après le big bang, où l'équilibre thermodynamique était rompu.
3. Un mécanisme, dans les interactions fondamentales, qui viole la conservation du nombre de baryons, c.-à-d., qui peut changer le nombre net de quarks.

La première condition est respectée par l'interaction faible, même si on ne connaît pas explicitement la cause profonde de cette brisure.⁸ La deuxième est acquise aussi, car il suffit que l'expansion de l'univers après le big bang soit, à un moment donné, plus rapide que le temps moyen d'une réaction importante. La troisième nécessite d'aller au-delà du modèle standard, vers un modèle d'unification des forces fondamentales.

Mésons B La brisure de *CP* n'est pas l'apanage des kaons, quoique les kaons neutres soient à ce jour le principal système physique où la brisure de *CP* (ou de T) a été observée. La violation de *CP* a été observée plusieurs fois depuis 2001 dans des expériences impliquant des mésons *B*. Les mésons B^0 et \bar{B}^0 sont formés d'un quark *b* et d'un antiquark *d*, ou l'inverse, et sont des cousins massifs des kaons. Les expériences dédiées à la production de mésons *B* sont appelées des « usines à *B* » (*B-factories*, en anglais). Une usine à *B* se base sur un collisionneur électron-positron. La masse invariante de la paire $e\bar{e}$ est ajustée juste à la valeur de la résonance $\Upsilon(4s)$, soit

8. On sait cependant comment la paramétrer : un élément complexe de la matrice CKM suffit. Mais cela n'explique pas comment une telle phase dans la matrice CKM est apparue.

un état excité du quarkonium $\bar{b}b$ ayant une énergie juste au-dessus du seuil de la désintégration $\Upsilon \rightarrow B\bar{B}$. La production des mésons B est alors abondante. Afin de produire un faisceau de mésons B , on donne aux faisceaux d'électrons et de positrons des énergies différentes (on dit que les deux faisceaux sont « asymétriques »), de sorte que le référentiel du centre de masse se déplace à une certaine vitesse, qui est grosso modo celle des mésons B produits.

G L'unification des forces

Le modèle standard, en dépit du remarquable schéma d'intégration des interactions fondamentales qu'il représente, ne les unifie pas vraiment en une seule théorie, car les trois interactions – forte, faible et électromagnétique – sont décrites par des constantes de couplages indépendantes. Certaines caractéristiques du modèle standard demandent d'ailleurs des explications qui ne sont pas intégrées au modèle lui-même :

1. Pourquoi y a-t-il trois familles de particules élémentaires?
2. Pourquoi la charge électrique est-elle quantifiée?
3. Quelle est l'origine des paramètres du modèle standard? Ces paramètres sont au nombre de 17 : les masses des particules (six pour les quarks, trois pour les leptons), les trois constantes de couplage, les quatre paramètres indépendants de la matrice CKM et enfin la valeur de ν dans le mécanisme de Higgs.

Depuis les années 1970, plusieurs théories ont été proposées dans le but de répondre à certaines de ces questions et d'unifier les interactions électrofaible et forte dans une même théorie de jauge. Aucune de ces théories n'a été retenue définitivement, car leurs prédictions sont très difficiles à vérifier. De telles théories portent le nom de *théories de grande unification*, ou TGU (en anglais *grand unified theories*, ou GUTs). L'idée de base est de supposer qu'un groupe de jauge supérieur G gouverne toutes les forces également. Les groupes de jauge de la QCD et de la théorie électrofaible sont des sous-groupes de G :

$$SU(3) \times SU(2) \times U(1) \subset G \quad (6.112)$$

La théorie de jauge basée sur G subit un mécanisme de brisure spontanée de la symétrie, semblable au mécanisme de Higgs, mais à une échelle d'énergie beaucoup plus élevée, appelée *échelle de grande unification*.

Quantification de la charge électrique Une telle théorie expliquerait, par exemple, la quantification de la charge électrique. Pour comprendre cette explication, procédons

encore une fois par analogie avec le moment cinétique. Considérons un atome plongé dans un champ magnétique, et en particulier un multiplet d'états de moment cinétique $j > 0$. Dans le champ magnétique, les $2j+1$ états du multiplet se séparent en niveaux d'énergie différents, avec des valeurs propres de J_z également espacées entre $-j$ et $+j$. Dans cette situation, le système ne jouit plus d'une invariance par rotation complète – par le groupe $SO(3) \sim SU(2)$ – mais d'une invariance réduite par le groupe $SO(2) = U(1)$, en raison de la présence du champ magnétique. Les représentations du groupe $U(1)$ sont toutes unidimensionnelles, car $U(1)$ est abélien, et sont donc spécifiées par une phase : $e^{im\varphi}$, où φ est l'angle de rotation par rapport à l'axe du

champ. D'un point de vue strictement mathématique, m peut prendre des valeurs quelconques si la symétrie $U(1)$ est le fin mot de l'histoire. Cependant, comme la symétrie $U(1)$ dans ce cas provient de la brisure d'une symétrie plus grande – en l'occurrence $SU(2)$ – les valeurs de m sont dictées par les représentations de ce groupe plus grand : elles vont de $-j$ à j et sont espacées de 1. En particulier, leur somme est nulle. Ce dernier point est une propriété générale de toute inclusion du groupe $U(1)$ dans un groupe de Lie plus grand.

Maintenant, portons notre attention sur les valeurs de l'hypercharge faible (Y) dans le modèle standard. On constate que la somme des valeurs de Y est nulle à l'intérieur d'une même famille (il faut compter la valeur de Y associée à chaque quark trois fois, en raison de la couleur). Ceci est considéré comme un signe sûr que la symétrie de jauge $U(1)_Y$ est en faite dérivée d'une symétrie plus grande, une symétrie de jauge nécessairement : les symétries de jauge imbriquées sont le reflet d'interactions imbriquées.

Unification des constantes de couplage L'unification des interactions sous un groupe de jauge unique a des implications sur la dépendance des constantes de couplage en fonction de l'énergie. La constante de couplage unique g_u de la théorie unifiée se sépare en trois constantes g_s , g et g' pour les parties forte et électrofaible de la théorie résultant de la brisure spontanée, et ces trois constantes de couplage évoluent différemment en fonction de l'échelle d'énergie et en accord avec l'éq. (5.153) :

$$\frac{1}{\alpha(s)} = \frac{1}{\alpha(s_0)} + b \ln \frac{s}{s_0}$$

où la constante b dépend de la théorie (N_g est le nombre de familles de particules élémentaires) :

$$\begin{aligned} 12\pi b &= 33 - 4N_g & SU(3) - \text{QCD} \\ 12\pi b &= 22 - 4N_g + \frac{1}{2} & SU(2) - \text{faible} \\ 12\pi b &= -4N_g & U(1) \end{aligned} \tag{6.113}$$

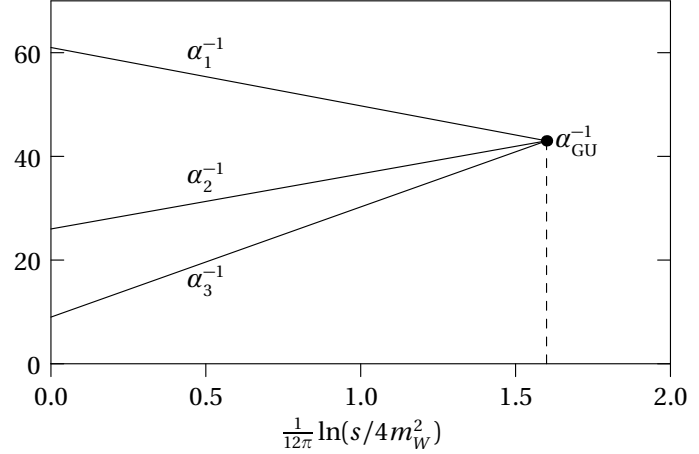
Les termes positifs de ces expressions proviennent de boucles de bosons (33 pour les bosons de jauge de $SU(3)$, 22 pour ceux de $SU(2)$, $\frac{1}{2}$ pour le boson de Higgs. Les termes négatifs proviennent des boucles de fermions. Les trois constantes de couplages sont connues à des précisions diverses. En particulier, on a mesuré la constante α_3 de la QCT avec la meilleure précision à l'échelle de la masse du méson Υ , et on trouve $\alpha_3(M_\Upsilon^2) \sim 0,16 \pm 0,02$. Pour que ces trois constantes coïncident à une échelle d'énergie $s = M_{\text{GU}}^2$ donnée, il faut que les conditions suivantes soient remplies (voir fig. 6.10) :

$$\sin^2 \theta_W = 0,21 \pm 0,01 \quad M_{\text{GU}} = 2 - 8 \times 10^{14} \text{ GeV} \quad \alpha_{\text{GU}} = 0,024 \tag{6.114}$$

où α_{GU} est la constante de couplage de la TGU à l'échelle d'énergie M_{GU} . Il est remarquable que l'angle de Weinberg prédit par l'hypothèse de la TGU concorde avec les valeurs mesurées autrement. Cependant l'échelle d'énergie M_{GU} à laquelle se produit cette unification est cent milliards de fois plus grande que celle que peut produire le plus puissant des accélérateurs présentement en fonction. C'est donc dire que la physique des interactions unifiées est loin de nos capacités expérimentales.

FIGURE 6.10

Évolution schématique, en fonction de l'échelle d'énergie, des constantes de couplages de la QCD (α_3) et de la théorie électrofaible (α_2 pour $SU(2)$ et α_1 pour $U(1)_Y$). Les trois constantes convergent à peu près à une échelle d'énergie appelée échelle de grande unification, et qui vaut environ 10^{14} GeV, après quoi une seule constante de couplage α_{GU} émerge.



Théorie unifiée basée sur le groupe $SU(5)$

Nous allons décrire brièvement la plus connue d'entre elles, basée sur le groupe de jauge $SU(5)$, proposée par S. GLASHOW et H. GEORGI.

Dans cette théorie, les quarks des trois couleurs et les leptons sont intégrés, famille par famille, dans des représentations du groupe $SU(5)$. La plus simple de ces représentations est la représentation fondamentale, de dimension 5, et on y agence les quarks d et les leptons comme suit :

$$\mathbf{5} : (d_1, d_2, d_3, e^c, -\nu^c)_R \quad \mathbf{5}^c : (d_1^c, d_2^c, d_3^c, e, -\nu)_L \quad (6.115)$$

où l'indice 1,2,3 des quarks désigne la couleur, et le symbole c désigne l'antiparticule (c.-à-d. $e^c = e^+$). $SU(5)$ possède aussi une représentation de dimension 10, sous la forme d'une matrice antisymétrique. Les fermions restants y sont agencés comme suit :

$$\mathbf{10} : \begin{pmatrix} 0 & u_3^c & -u_2^c & u_1 & d_1 \\ & 0 & u_1^c & u_2 & d_2 \\ & & 0 & u_3 & d_3 \\ & & & 0 & e^c \\ & & & & 0 \end{pmatrix}_L \quad (6.116)$$

Notons que $(u_i^c)_L$ contient en fait u_{iR} , et donc le contenu de $\mathbf{10}$ et $\mathbf{5}$ est exhaustif.

Les champs de jauge de la théorie $SU(5)$ sont au nombre de $5^2 - 1 = 24$. De ce nombre, huit sont les précurseurs des gluons et quatre les précurseurs des bosons W^\pm , Z et du photon. Il reste douze bosons de jauge qui acquièrent une masse de l'ordre de 10^{14} GeV lors de la brisure spontanée de $SU(5)$. Ces douze bosons de jauge sont notés X_a^\pm et Y_a^\pm et forment des doublets d'isospin faible (X, Y) et des triplets de couleur (X_1, X_2, X_3) et (Y_1, Y_2, Y_3). Les bosons X portent en outre une charge électrique $\pm \frac{4}{3}$ et les bosons Y une charge $\pm \frac{1}{3}$.

La théorie $SU(5)$ explique naturellement que la charge des quarks doit être en multiples de $\frac{1}{3}$ et celle des leptons en multiples de 1. Une autre caractéristique est que cette théorie fournit un mécanisme pour la non-conservation du nombre baryonique. Autrement dit, il existe des processus qui permettent la transformation d'un quark en lepton, par émission ou absorption d'un boson de jauge X . Les diagrammes de Feynman typiques impliquant les bosons X et Y sont illustrés à la fig. 6.11.

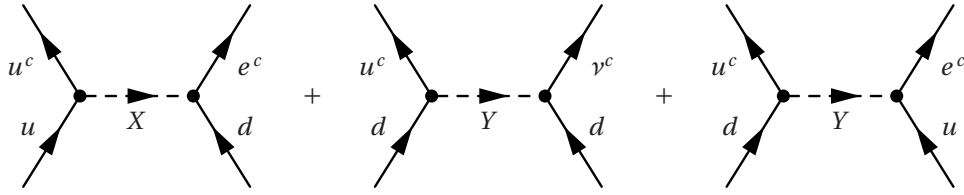


FIGURE 6.11

Diagrammes de Feynman impliquant les bosons X et Y de la théorie de grande unification $SU(5)$. Notez qu'une antiparticule entrante, comme u^c , équivaut à une particule sortante, de sorte que les processus illustrés ici ne conservent pas la nombre baryonique (ou le nombre de quarks).

En particulier, un processus possible est la désintégration du proton, qui pourrait s'effectuer par l'intermédiaire d'un boson X , comme illustré ci-dessous. Un d spectateur est sous-entendu, et le proton se désintègre en un π^0 et un positron :



La vie moyenne estimée du proton par ce processus est située entre 10^{28} et 10^{32} années. L'estimation expérimentale de la vie moyenne du proton dans le processus $p \rightarrow e^+ \pi^0$ est supérieure à une valeur allant de 10^{31} et 10^{33} années. La prédiction de la théorie $SU(5)$ est donc marginalement rejetée par l'expérience. Mais les incertitudes sont énormes tant que la masse des hypothétiques bosons X et Y n'est pas estimée avec plus de précision.

Autres théories unifiées Une autre théorie d'unification intéressante est basée sur le groupe $SO(10)$. Elle a l'avantage esthétique que tous les fermions d'une même famille sont groupés en une seule représentation de dimension 16 (dite spinorielle) de $SO(10)$. Le fermion supplémentaire représente alors le neutrino droit. Un mécanisme élégant de brisure de symétrie, dit « en dents de scie » permet naturellement d'expliquer que les neutrinos ont une masse beaucoup plus petite que les autres fermions. Le groupe $SO(10)$ est de dimension 45, ce qui suppose l'existence d'un plus grand nombre encore de champs de jauge massifs à l'échelle M_{GU} .

H Problèmes

Question 6.1

Pourquoi un terme de masse pour les leptons ou les quarks viole-t-il la symétrie de jauge $SU(2)$?

Question 6.2

Neutrinos droits Si les neutrinos droits (ν_R) existaient en tant que singulets de $SU(2)$, quel devrait être leur hypercharge faible Y ? Pourrait-on alors écrire un couplage de Yukawa avec les leptons gauches et le champ Φ menant à un terme de masse pour les neutrinos ? Quelle serait, par contre, la conséquence d'un tel terme ?

Problème 6.1 Théorie de Fermi de l'interaction faible

Dans la théorie de Fermi des interactions faibles, le hamiltonien d'interaction est donné par l'expression suivante :

$$H_I = -\frac{G_F}{\sqrt{2}} \int d^3r (\bar{p} \gamma^\mu n) (\bar{e} \gamma_\mu \nu) \quad (6.118)$$

où p , n , ν , et e sont les spineurs décrivant respectivement le proton, le neutron, le neutrino et l'électron. Par définition, $\gamma_\mu = g_{\mu\nu} \gamma^\nu$.

Considérez que l'état initial $|i\rangle$ du système est un neutron au repos, et que l'état final $|f\rangle$ contient un proton, un électron et un antineutrino. Supposons que le proton, l'électron et l'antineutrino sont respectivement dans des états de spineur $u_p(\mathbf{p}_p)$, $u_e(\mathbf{p}_e)$ et $\nu_\nu(\mathbf{p}_\nu)$, et que le neutron d'origine a un spin 'up'.

Utilisez la règle d'or pour calculer l'amplitude associée à ce processus de désintégration.

Problème 6.2 Particule de Higgs

Écrivons le champ de Higgs Φ comme la somme

$$\Phi = \frac{1}{\sqrt{2}} \begin{pmatrix} 0 \\ \nu + h \end{pmatrix}$$

où h est une déviation par rapport à la valeur de Φ dans l'état fondamental. Le champ h décrit justement ce qu'on appelle le boson de Higgs.

A En appliquant cette définition au Lagrangien du modèle standard, écrivez les termes impliquant h . Exprimez vos résultats en fonction des champs W^\pm et Z . Quelle est la masse de la particule décrite par le champ h , en fonction des paramètres à votre disposition ? Illustrez les vertex d'interaction impliquant le champ h et les champs de jauge.

B Comment la particule de Higgs, décrite par le champ h , interagit-elle avec les fermions du modèle standard ? Écrivez le Lagrangien d'interaction. Donnez un exemple de processus par lesquels la particule de Higgs pourrait se désintégrer.

Problème 6.3 Processus de désintégration

À l'aide d'un diagramme de flot de particules (quarks ou autres), illustrez comment peuvent se désintégrer les particules suivantes : essayez de donner plusieurs possibilités là où c'est possible, mais concentrez-vous sur les processus importants. Consultez le [données disponibles](#) pour vérifier que les désintégrations peuvent se produire comme vous le suggérez, et notez le rapport d'embranchement ainsi que la vie moyenne.

A Le neutron

B Le muon

C Le lepton tau (τ)

D Le pion π^+

E Le pion neutre π^0

F Le méson lourd Υ ($b\bar{b}$)

Problème 6.4 Désintégration des kaons

Dessinez des diagrammes de Feynman illustrant les processus de désintégration suivants :

A $K^+ \rightarrow \mu^+ + \nu_\mu$

B $K^0 \rightarrow \pi^+ + \pi^-$

Problème 6.5 Désintégration du pion

A Démontrez l'équation (6.97).

B Démontrez ensuite l'équation (6.99).

C Démontrez enfin l'équation (6.100).

CHAPITRE 7

ANNEXES

A Constantes physiques et facteurs de conversion

constante de Planck	\hbar	$6,582\,118\,99(16) \times 10^{-22} \text{ MeV.s}$ $6,582\,118\,99(16) \times 10^{-7} \text{ eV.ns}$
	$\hbar c$	$197,326\,9631(49) \text{ MeV.fm}$
constante de structure fine	α	$1/137,035\,999\,679(94)$
unité de masse atomique	u	$931,494\,028(23) \text{ MeV}/c^2$
nombre d'Avogadro	N_0	$6,022\,141\,79(30) \times 10^{23} \text{ mol}^{-1}$
constante de Boltzmann	k_B	$8,617\,343(15) \times 10^{-5} \text{ eV/K}$

B Unités électromagnétiques

Les unités électromagnétiques sont souvent l'objet de confusion, car plusieurs systèmes différents sont utilisés en physique. Le tableau ci-dessous résume les différents systèmes (nous avons restauré la vitesse de la lumière c , pour rendre la discussion indépendante du système d'unités naturelles). On y définit plusieurs constantes, en fonction desquelles les équations de Maxwell microscopiques prennent la forme suivante :

$$\begin{aligned} \nabla \cdot \mathbf{E} &= 4\pi k_c \rho & \nabla \cdot \mathbf{B} &= 0 \\ \nabla \wedge \mathbf{E} + \alpha_L \frac{\partial \mathbf{B}}{\partial t} &= 0 & \nabla \wedge \mathbf{B} - \frac{\alpha_B}{k_c} \frac{\partial \mathbf{E}}{\partial t} &= 4\pi \alpha_B \mathbf{J} \end{aligned} \quad (7.1)$$

D'autres parts, la loi de Coulomb et la loi de Biot-Savart prennent les formes suivantes :

$$\mathbf{F} = k_c \frac{q_1 q_2}{r^2} \quad \mathbf{dB} = \alpha_B I \frac{d\mathbf{l} \wedge \mathbf{r}}{r^3} \quad (7.2)$$

système	k_c	α_B	ϵ_0	μ_0	$k_a = \frac{k_c}{c^2}$	$\alpha_L = \frac{k_c}{\alpha_B c^2}$	$\lambda = 4\pi k_c \cdot \epsilon_0$	λ'
ESU (ou stat-)	1	c^{-2}	1	c^{-2}	c^{-2}	1	4π	4π
EMU (ou ab-)	c^2	1	c^{-2}	1	1	1	4π	4π
gaussien	1	c^{-1}	1	1	c^{-2}	c^{-1}	4π	4π
Heaviside-Lorentz	$\frac{1}{4\pi}$	$\frac{1}{4\pi c}$	1	1	$\frac{1}{4\pi c^2}$	c^{-1}	1	1
SI	$\frac{1}{4\pi\epsilon_0}$	$\frac{\mu_0}{4\pi}$	ϵ_0	μ_0	μ_0	1	1	1

TABLE 7.1
Constantes électromagnétiques dans les différents systèmes d'unités en usage

La force par unité de longueur entre deux fils parallèles séparés d'une distance d et portant des courants I_1 et I_2 est

$$\frac{dF}{dL} = 2k_a \frac{I_1 I_2}{d} \quad (7.3)$$

et la force magnétique s'exprime comme

$$\mathbf{F} = \alpha_L q \mathbf{v} \wedge \mathbf{B} \quad (7.4)$$

Les champs macroscopiques \mathbf{H} et \mathbf{D} sont définis de la manière suivante :

$$\mathbf{D} = \epsilon_0 \mathbf{E} + \lambda \mathbf{P} \quad \mathbf{H} = \frac{1}{\mu_0} \mathbf{B} - \lambda' \mathbf{M} \quad (7.5)$$

De sorte que les équations de Maxwell pour les champs macroscopiques prennent la forme

$$\begin{aligned} \nabla \cdot \mathbf{D} &= 4\pi \epsilon_0 k_c \rho & \nabla \cdot \mathbf{B} &= 0 \\ \nabla \wedge \mathbf{E} + \alpha_L \frac{\partial \mathbf{B}}{\partial t} &= 0 & \nabla \wedge \mathbf{H} - \frac{\alpha_B}{k_c \mu_0 \epsilon_0} \frac{\partial \mathbf{D}}{\partial t} &= \frac{4\pi \alpha_B}{\mu_0} \mathbf{J} \end{aligned} \quad (7.6)$$

Un système dans lequel les facteurs géométriques liés à la dimension de l'espace (comme 4π) n'apparaissent pas dans les équations de Maxwell est qualifié de *rationalisé*. Seuls le système SI et celui de Heaviside-Lorentz le sont.

Le système de Heaviside-Lorentz est celui qui combine le plus d'avantages fondamentaux :

1. Il est rationalisé, de sorte que les équations fondamentales sont indépendantes de la dimension de l'espace.
2. Tous les champs ont les mêmes unités, car $\epsilon_0 = \mu_0 = \lambda = \lambda' = 1$.
3. La vitesse de la lumière y apparaît explicitement et aucune constante artificielle ne figure dans les équations fondamentales.

Formule	Heaviside-Lorentz	gaussien	SI
Force de Lorentz	$\mathbf{F} = q \left(\mathbf{E} + \frac{1}{c} \mathbf{v} \wedge \mathbf{B} \right)$	$\mathbf{F} = q \left(\mathbf{E} + \frac{1}{c} \mathbf{v} \wedge \mathbf{B} \right)$	$\mathbf{F} = q (\mathbf{E} + \mathbf{v} \wedge \mathbf{B})$
Lois de Coulomb	$F = \frac{1}{4\pi} \frac{q_1 q_2}{r^2}$	$F = \frac{q_1 q_2}{r^2}$	$F = \frac{1}{4\pi\epsilon_0} \frac{q_1 q_2}{r^2}$
Loi de Biot & Savart	$\mathbf{B} = \frac{I}{4\pi c} \oint \frac{d\mathbf{l} \wedge \mathbf{e}_r}{r^2}$	$\mathbf{B} = \frac{I}{c} \oint \frac{d\mathbf{l} \wedge \mathbf{e}_r}{r^2}$	$\mathbf{B} = \frac{\mu_0 I}{4\pi} \oint \frac{d\mathbf{l} \wedge \mathbf{e}_r}{r^2}$
induction électrique	$\mathbf{D} = \mathbf{E} + \mathbf{P}$	$\mathbf{D} = \mathbf{E} + 4\pi\mathbf{P}$	$\mathbf{D} = \epsilon_0 \mathbf{E} + \mathbf{P}$
induction magnétique	$\mathbf{B} = \mathbf{H} + \mathbf{M}$	$\mathbf{B} = \mathbf{H} + 4\pi\mathbf{M}$	$\mathbf{B} = \mu_0 (\mathbf{H} + \mathbf{M})$
loi de Gauss	$\nabla \cdot \mathbf{E} = \rho$	$\nabla \cdot \mathbf{E} = 4\pi\rho$	$\nabla \cdot \mathbf{E} = \frac{\rho}{\epsilon_0}$
loi de Faraday	$\nabla \wedge \mathbf{E} = -\frac{1}{c} \frac{\partial \mathbf{B}}{\partial t}$	$\nabla \wedge \mathbf{E} = -\frac{1}{c} \frac{\partial \mathbf{B}}{\partial t}$	$\nabla \wedge \mathbf{E} = -\frac{\partial \mathbf{B}}{\partial t}$
loi d'Ampère	$\nabla \wedge \mathbf{B} - \frac{1}{c} \frac{\partial \mathbf{E}}{\partial t} = \frac{1}{c} \mathbf{J}$	$\nabla \wedge \mathbf{B} - \frac{1}{c} \frac{\partial \mathbf{E}}{\partial t} = \frac{4\pi}{c} \mathbf{J}$	$\nabla \wedge \mathbf{B} - \epsilon_0 \mu_0 \frac{\partial \mathbf{E}}{\partial t} = \mu_0 \mathbf{J}$

TABLE 7.2

Relations électromagnétiques dans les trois principaux systèmes d'unités.

C Vecteurs et tenseurs

L'objectif de cette section est d'introduire les notions de composantes covariantes et contravariantes de vecteurs et de tenseurs. Pour rendre la discussion plus intuitive, nous travaillerons dans l'espace cartésien à trois dimensions, mais les concepts seront immédiatement généralisables à l'espace-temps quadridimensionnel.

Composantes covariantes et contravariantes

Considérons une base $\{\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3\}$ dans l'espace tridimensionnel. Cette base n'est pas nécessairement orthonormée. Tout vecteur \mathbf{A} peut être décomposé selon cette base :

$$\mathbf{A} = A^1 \mathbf{e}_1 + A^2 \mathbf{e}_2 + A^3 \mathbf{e}_3 = A^i \mathbf{e}_i \quad (7.7)$$

où on applique la convention de sommation sur les indices répétés. Les trois quantités A^i sont appelées *composantes contravariantes* du vecteur \mathbf{A} . La raison de cette appellation est donnée dans ce qui suit, et est au coeur de la présente annexe. Supposons qu'on procède à un changement de base, c'est-à-dire qu'on définit une nouvelle base \mathbf{e}'_i :

$$\mathbf{e}'_i = S_i^j \mathbf{e}_j \quad (7.8)$$

où la matrice S est non singulière (indice de rangée à gauche, indice de colonne à droite). Le vecteur s'exprime également dans l'une ou l'autre base :

$$\mathbf{A} = A^i \mathbf{e}_i = A'^i \mathbf{e}'_i = A'^i S_i^j \mathbf{e}_j \quad (7.9)$$

Donc $A^j = A'^i S_i^j$. Les composantes A^j doivent s'exprimer en fonction des composantes A'^i par l'intermédiaire d'une matrice \tilde{S} :

$$A'^j = \tilde{S}^j_i A^i \quad (7.10)$$

qui n'est autre que la transposée de l'inverse de S :

$$\tilde{S}^j_i S_k^i = \delta_k^j \quad \text{ou} \quad \tilde{S} S^T = 1 \quad (7.11)$$

Autrement dit, les composantes contravariantes se transforment à l'aide de la matrice inverse transposée de celle utilisée pour transformer les vecteurs de base (d'où l'expression *contravariante*).

À la base $\{\mathbf{e}_i\}$ on associe une *base duale* $\{\mathbf{e}^i\}$ (l'indice est maintenant en haut) définie par la relation

$$\mathbf{e}^i \cdot \mathbf{e}_j = \delta_j^i \quad (7.12)$$

Comme la relation entre une base et sa duale est la même pour toutes les bases ($\mathbf{e}^i \cdot \mathbf{e}_j = \mathbf{e}'^i \cdot \mathbf{e}'_j = \delta_j^i$) et qu'elle a la même structure que la relation (7.9), on en déduit immédiatement que les vecteurs de la base duale se transforment aussi comme les composantes contravariantes lors d'un changement de base :

$$\mathbf{e}'^j = \tilde{S}^j_i \mathbf{e}^i \quad (7.13)$$

Autrement dit, la base duale est contravariante.

Tout vecteur peut également être représenté de manière unique sur la base duale, mais cette fois par des composantes dites *covariantes* :

$$\mathbf{A} = A_i \mathbf{e}^i = A'_i \mathbf{e}'^i \quad (7.14)$$

Il est maintenant évident que les composantes covariantes méritent ce nom parce qu'elles se transforment de la même manière que les vecteurs de base :

$$A'_i = S_i^j A_j \quad (7.15)$$

Les composantes covariantes s'obtiennent par projection du vecteur sur la base originale :

$$A_i = \mathbf{A} \cdot \mathbf{e}_i \quad (7.16)$$

Tenseurs Si $\{\mathbf{e}_i\}$ est une base de l'espace vectoriel V , alors les produits $\{\mathbf{e}_i \otimes \mathbf{e}_j\}$ forment une base de l'espace produit tensoriel $V \otimes V$. on peut aussi utiliser la base duale, ou deux bases mixtes : $\{\mathbf{e}^i \otimes \mathbf{e}^j\}$, $\{\mathbf{e}_i \otimes \mathbf{e}^j\}$ et $\{\mathbf{e}^i \otimes \mathbf{e}_j\}$. Dans cet espace produit, on définit des tenseurs de rang 2, caractérisés par des composantes à deux indices :

$$\mathbf{T} = T^{ij} \mathbf{e}_i \otimes \mathbf{e}_j = T_{ij} \mathbf{e}^i \otimes \mathbf{e}^j = T^i_j \mathbf{e}_i \otimes \mathbf{e}^j = T_i^j \mathbf{e}^i \otimes \mathbf{e}_j \quad (7.17)$$

On constate immédiatement que, lors d'un changement de base, ces composantes se transforment de la manière suivante :

$$\begin{aligned} T'_{ij} &= S_i^k S_j^l T_{kl} \\ T'^{ij} &= \tilde{S}^i_k \tilde{S}^j_l T^{kl} \\ T'^i_j &= \tilde{S}^i_k S_j^l T^k_l \\ T_i'^j &= S_i^k \tilde{S}^j_l T_k^l \end{aligned} \quad (7.18)$$

Par abus de langage, on donne le nom de tenseur aux composantes elles-mêmes, et on appelle T_{ij} un tenseur covariant, T^{ij} un tenseur contravariant, alors que les ensembles T_i^j et T^i_j sont qualifiés de *tenseurs mixtes*.

On définit généralement des tenseurs de rang N , sur l'espace produit tensoriel $V \otimes V \otimes \dots \otimes V$ (N facteurs), qui comportent N indices. Un tenseur de rang un est un vecteur. Un tenseur de rang zéro est un scalaire (une quantité invariante, c'est-à-dire qui est indépendante de la base utilisée).

Tenseur métrique La base $\{\mathbf{e}_i\}$ peut bien sûr être exprimée dans la base duale :

$$\mathbf{e}_i = g_{ij} \mathbf{e}^j \quad (7.19)$$

Les coefficients g_{ij} sont également les produits scalaires des vecteurs de base :

$$\mathbf{e}_i \cdot \mathbf{e}_j = g_{ik} \mathbf{e}^k \cdot \mathbf{e}_j = g_{ik} \delta_j^k = g_{ij} \quad (7.20)$$

La relation inverse s'exprime ainsi :

$$\mathbf{e}^i = g^{ij} \mathbf{e}_j \quad \text{où} \quad g^{ik} g_{kj} = \delta_j^i \quad \text{et} \quad \mathbf{e}^i \cdot \mathbf{e}^j = g^{ij} \quad (7.21)$$

Comme les vecteurs de base \mathbf{e}_i et \mathbf{e}^i sont respectivement covariants et contravariants, les quantités g_{ij} et g^{ij} se transforment bien comme les composantes covariantes et contravariantes d'un tenseur de rang 2. On l'appelle le *tenseur métrique*.

Le tenseur métrique peut être utilisé pour passer d'une composante covariante à contravariante et vice-versa. Ceci est vrai (par définition) pour les vecteurs de base, mais aussi pour les composantes. Explicitement :

$$\begin{aligned} \mathbf{A} &= A^i \mathbf{e}_i = A^i g_{ij} \mathbf{e}^j \implies A_j = g_{ij} A^i \\ \mathbf{A} &= A_i \mathbf{e}^i = A_i g^{ij} \mathbf{e}_j \implies A^j = g^{ij} A_i \end{aligned} \quad (7.22)$$

Plus généralement, le tenseur métrique peut être appliqué de cette manière à tout indice, covariant ou contravariant :

$$T_{ij} = g_{ik} g_{jl} T^{kl} \quad , \quad T_i^j = g^{jk} T_{ik} \quad , \text{ etc.} \quad (7.23)$$

Contraction On appelle contraction d'un indice covariant et d'un indice contravariant la somme sur les valeurs égales de ces indices, comme dans un produit scalaire. Ainsi, dans l'expression $A^i \mathbf{e}_i$, on dit que les deux indices sont contractés. Autre exemple : le produit scalaire de deux vecteurs \mathbf{A} et \mathbf{B} s'exprime ainsi :

$$\mathbf{A} \cdot \mathbf{B} = (A^i \mathbf{e}_i) \cdot (B_j \mathbf{e}^j) = A^i B_j \delta_i^j = A^i B_i \quad (7.24)$$

Les indices d'un tenseur de rang deux peuvent aussi être contractés : T^i_i .

En général, la contraction de deux indices mène soit à une quantité invariante, soit à une quantité de rang inférieur de deux à la quantité originale. Par exemple, d'un tenseur de rang 4 R_{iklm} (comme le tenseur de courbure en géométrie riemannienne), on peut former un tenseur de rang deux par contraction : $R_{ik} = R^l_{ilk} = g^{ml} R_{milk}$.

D Groupes définis par des matrices

Un ensemble G forme un *groupe* si une loi de composition (appelons-la multiplication) a été définie sur cet ensemble, telle que

1. L'ensemble est fermé par cette loi : si $a \in G$ et $b \in G$, alors $ab \in G$. Dans ce qui suit, cette condition sera souvent appelée la *règle de groupe*.
2. La loi est associative : $(ab)c = a(bc)$.
3. Il existe un élément neutre e , tel que $ea = ae = a$, pour tout élément a de G .
4. Tout élément a possède un inverse a^{-1} dans G tel que $a^{-1}a = e$.

Si la multiplication est commutative ($ab = ba$), le groupe est dit *abélien*. Dans le cas contraire, plus intéressant de notre point de vue, le groupe est dit *non abélien*. Un sous-ensemble H de G est appelé *sous-groupe* de G s'il est lui-même un groupe, avec la même loi de multiplication que G .

Groupes finis, discrets et infinis Un groupe est dit *fini* s'il compte un nombre fini d'éléments. Il est *infini* dans le cas contraire. Il est dit *discret* si ses éléments sont clairement séparés, plus précisément s'ils sont dénombrables. Il est continu si ses éléments forment un espace dans lequel la notion de continuité est définie. Plus précisément, un *groupe de Lie* est un groupe dont les éléments peuvent être repérés par un système de coordonnées tiré de \mathbb{R}^n (on dit alors que la *dimension* du groupe est n , le nombre de paramètres nécessaire pour spécifier un élément du groupe).

Groupes définis par des matrices Il est plus simple à ce stade-ci de procéder directement à la définition des groupes de Lie qui nous intéresseront par la suite. Ceux-ci peuvent être définis par des matrices. La loi de multiplication du groupe est alors la loi de multiplication des matrices (réelles ou complexes). Les matrices formant le groupe doivent toutes être régulières (c'est-à-dire avoir un inverse). Le groupe général de toutes les matrices régulières d'ordre n est appelé $GL(n)$. Tous les groupes de Lie « classiques » sont des sous-groupes de $GL(n)$. Il est évident que $GL(n)$ est un groupe : la multiplication des matrices est associative et chaque matrice de $GL(n)$ possède un inverse par définition. D'autre part, le produit de deux matrices régulières est encore une matrice régulière. Les groupes matriciels les plus importants sont :

1. Le *groupe unitaire* $U(n)$, formé des matrices complexes unitaires d'ordre n . Bien entendu, le produit de deux matrices unitaires est encore unitaire et la règle de groupe est respectée. Le nombre de paramètres nécessaires pour spécifier une matrice complexe d'ordre n est $2n^2$. L'unitarité $U^\dagger U = 1$ impose $n + 2\frac{1}{2}n(n-1)$ conditions supplémentaires, ce qui réduit le nombre de paramètres à n^2 . La dimension du groupe $U(n)$ est donc $d = n^2$.
2. Son sous-groupe $SU(n)$, restreint par la condition $\det U = 1$. Cette condition respecte la règle de groupe. Une condition supplémentaire est imposée et le nombre de paramètres est réduit à $d = n^2 - 1$.
3. Le *groupe orthogonal* $O(n)$, formé des matrices réelles orthogonales d'ordre n . Ici encore la règle de groupe est respectée, car le produit de deux matrices orthogonales est aussi une matrice orthogonale. Le nombre de paramètres est n^2 , moins le nombre de contraintes indépendantes imposées par l'orthogonalité, soit $\frac{1}{2}n(n+1)$. Il reste $d = \frac{1}{2}n(n-1)$ paramètres.
4. Son sous-groupe $SO(n)$, restreint par la condition $\det O = 1$. Notons que cette condition ne change pas la dimension du groupe, car de toute manière $\det O = \pm 1$ dans $O(n)$. Topologiquement, elle restreint le groupe à une composante *connexe*. Autrement dit, tous les éléments de $SO(n)$ peuvent être obtenus en procédant de manière continue à partir de l'identité, ce qui n'est pas le cas de $O(n)$, puisque les matrices dites *impropres* ($\det O = -1$) produisent une inversion de l'espace à n dimensions sur lequel elles agissent. Le groupe des rotations en trois dimensions est précisément $SO(3)$.

Symbole	\mathbb{R} ou \mathbb{C}	Définition	Dimension	Rang
$U(n)$	\mathbb{C}	$U^\dagger U = 1$	n^2	n
$SU(n)$	\mathbb{C}	$U^\dagger U = 1, \det U = 1$	$n^2 - 1$	$n - 1$
$O(n)$	\mathbb{R}	$O^T O = 1$	$n(n-1)/2$	$[n/2]$
$SO(n)$	\mathbb{R}	$O^T O = 1, \det O = 1$	$n(n-1)/2$	$[n/2]$

Remarques :

- ♦ $SO(n)$ est bien sûr un sous-groupe de $SU(n)$.
- ♦ $SU(n-1)$ est un sous-groupe de $SU(n)$, comme $SO(n-1)$ de $SO(n)$ et ainsi de suite.

Générateurs de $SU(n)$ Toute matrice U de $SU(n)$ peut être représentée comme une exponentielle $U = e^{iH}$, où H est une matrice hermitienne de trace nulle. En effet, la contrainte d'unitarité de U signifie que

$$U^\dagger = U^{-1} \implies e^{-iH^\dagger} = e^{-iH} \implies H^\dagger = H \quad (7.25)$$

De plus, la contrainte $\det U = 1$ se traduit par $\text{Tr } H = 0$, en vertu de l'identité matricielle

$$\ln \det M = \text{Tr } \ln M \quad (7.26)$$

pour une matrice M quelconque.

Soit T^a ($a = 1, 2, \dots, d$) une base de matrices hermitiennes de trace nulle. Ici d est le nombre de matrices de ce type qui sont linéairement indépendantes. Pour le groupe $SU(n)$, $d = n^2 - 1$, soit

$d = 3$ pour $SU(2)$ et $d = 8$ pour $SU(3)$. Toute matrice hermitienne de trace nulle peut alors s'écrire comme $H = \omega_a T_a$, où les ω_a sont des paramètres réels. Enfin, toute matrice de $SU(n)$ peut alors s'écrire comme $U = e^{i\omega_a T_a}$. On appelle les matrices T_a les *générateurs* du groupe $SU(n)$, car elles peuvent générer tous les éléments du groupe, à partir de transformations proches de l'identité. Les T_a sont en fait une généralisation à $SU(n)$ des matrices du moment cinétique pour le groupe $SU(2)$.

Les générateurs de $SU(3)$ sont communément définis en fonction des matrices de Gell-Mann, sorte de généralisation des matrices de Pauli à l'ordre 3 :

$$\begin{aligned} \lambda_1 &= \begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} & \lambda_2 &= \begin{pmatrix} 0 & -i & 0 \\ i & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} & \lambda_3 &= \begin{pmatrix} 1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 0 \end{pmatrix} & \lambda_4 &= \begin{pmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ 1 & 0 & 0 \end{pmatrix} \\ \lambda_5 &= \begin{pmatrix} 0 & 0 & -i \\ 0 & 0 & 0 \\ i & 0 & 0 \end{pmatrix} & \lambda_6 &= \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{pmatrix} & \lambda_7 &= \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & -i \\ 0 & i & 0 \end{pmatrix} & \lambda_8 &= \frac{1}{\sqrt{3}} \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -2 \end{pmatrix} \end{aligned} \quad (7.27)$$

Les matrices de Gell-Mann sont normalisées afin que $\text{tr}(\lambda_a \lambda_b) = 2\delta_{ab}$. Les générateurs de $SU(3)$ sont généralement choisis comme $T_a = \frac{1}{2}\lambda_a$, de sorte que $\text{tr}(T_a T_b) = \frac{1}{2}\delta_{ab}$. Notons que deux des générateurs, T_3 et T_8 , commutent entre eux. Le nombre maximal de générateurs qui commutent au sein d'un groupe de Lie est appelé *rang* et apparaît dans le tableau ci-haut.

Constantes de structure Les générateurs de $SU(n)$ obéissent à des relations de commutation qui sont en quelque sorte une généralisation de celles qui s'appliquent aux composantes du moment cinétique :

$$[T_a, T_b] = i f_{abc} T_c \quad (7.28)$$

où les f_{abc} sont appelées *constantes de structure*. En effet, comme le commutateur de deux matrices hermitiennes est antihermitien et que la trace d'un commutateur de matrices finies s'annule, il s'ensuit que le commutateur $[T_a, T_b]$ est nécessairement i fois une matrice hermitienne sans trace, donc une combinaison des T_a . Les constantes de structure sont complètement antisymétriques en (a, b, c) ; démontrons cette assertion. Comme $\text{tr}(T_a T_b) = \frac{1}{2}\delta_{ab}$, on a $\text{tr}([T_a, T_b] T_c) = \frac{1}{2} i f_{abc}$. Or cette expression est manifestement antisymétrique en (a, b) , en raison du commutateur. En vertu de la propriété cyclique de la trace,

$$\text{tr}([T_a, T_b] T_c) = \text{tr}(T_a T_b T_c - T_b T_a T_c) = \text{tr}(T_b T_c T_a - T_c T_b T_a) = \text{tr}([T_b, T_c] T_a) \quad (7.29)$$

et donc f_{abc} est aussi antisymétrique en (b, c) . Il s'ensuit alors qu'il est antisymétrique en (a, c) et, généralement, par rapport à toute permutation des indices a, b, c . Ces propriétés sont vraies pour $SU(n)$ en général. Dans le cas de $SU(3)$, on a les valeurs particulières suivantes :

$$\begin{aligned} f_{123} &= 1 \\ f_{147} &= f_{246} = f_{257} = f_{345} = f_{516} = f_{637} = \frac{1}{2} \\ f_{458} &= f_{678} = \frac{\sqrt{3}}{2} \end{aligned} \quad (7.30)$$

E Tableau des principales particules subatomiques

symbole	nom	spin	masse (MeV)	composition
e^-	électron	1/2	0,511	—
μ^-	muon	1/2	105,65	—
τ^-	tauon	1/2	1777	—
ν_e, ν_μ, ν_τ	neutrinos	1/2	$> 2 \times 10^{-8}, < 10^{-6}$	—
p^+	proton	1/2	938,27	uud
n	neutron	1/2	939,56	udd
$\Delta^{++}, \Delta^+, \Delta^0, \Delta^-$		3/2	1232–1233	uuu, uud, udd, ddd
$\Sigma^+, \Sigma^0, \Sigma^-$		1/2	1189	uus, uds, dds
Λ		1/2	1115,68	uds
Ξ^0	Cascade	1/2	1314,86	uss
Ξ^-	Cascade	1/2	1321,71	dss
Ω^-		3/2	1672,45	sss
π^\pm	pion	0	139,57	$u\bar{d}, d\bar{u}$
π^0	pion	0	134,98	$u\bar{u}, d\bar{d}$
η		0	547,85	$u\bar{u}, d\bar{d}, s\bar{s}$
K^\pm	kaon	0	493,7	$u\bar{s}, s\bar{u}$
K^0, \bar{K}^0	kaon	0	497,6	$d\bar{s}, s\bar{d}$
D^\pm		0	1869,6	$d\bar{c}, c\bar{d}$
D^0, \bar{D}^0		0	1864,8	$u\bar{c}, c\bar{u}$
B^\pm		0	5279	$u\bar{b}, b\bar{u}$
B^0, \bar{B}^0		0	5279	$d\bar{b}, b\bar{d}$
η_c		0	2980	$c\bar{c}$
J/ψ		1	3097	$c\bar{c}$
$\Upsilon(1S)$		1	9460	$b\bar{b}$
γ	photon	1	0	—
W		1	80 385	—
Z		1	91 188	—
H^0	Higgs	0	125 090	—